

APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hotte, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 1213
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-79

alignment_scores:
Quality: 141.00 Length: 577
Ratio: 0.613 Gaps: 29
Percent Similarity: 39.861 Percent Identity: 24.957

alignment_block:
US-09-303-518D-125 x US-09-413-814-79 ..

Align seg 1/1 to: US-09-413-814-79 from: 1 to: 1213

```
21 TCTAAACCTGCCATCGCGG.....CAGAC 46
   ::::::::::::::::::::
15 AlaArpProAlaArgArgGlyAspArgArgGlyLeuAlaLeuArgSe 31
   ::::::::::::::::::::
47 CGGACCAACCGCTTACGACGGCCGCGCAT..... 77
   ::::::::::::::::::::
31 rAspArgGluArgAlaArgArgArgGlyHisGlyLeuGlyLeuGlyAlaG 48
   ::::::::::::::::::::
78 .....TACCGAGTCGCGCTGTCGCGAGAGATATCGCGGTATGCGGCC 122
   ::::::::::::::::::::
48 lyGlyArgArgAlaArgArgAlaAlaArgArgArgLeuProAlaPro 64
   ::::::::::::::::::::
122 .....
65 GlyProSerProArgArgGlyAlaProGlnAlaProProGlyAlaArgAr 81
   ::::::::::::::::::::
123 .....CTGATGAA.....A 132
   ::::::::::::::::::::
81 gGlyProProArgProHisAlaValGlyGlyArgArgAspAspArgLeuA 98
   ::::::::::::::::::::
133 GTCAAGAGAGCGGATGTC...CGTCAAAAAGAGCCCAAGTGTCTTTGAAGA 179
   ::::::::::::::::::::
98 laArgGlyHisArgAlaProArgArgArgArgGlyArgAlaValAlaArg 114
   ::::::::::::::::::::
180 .....CAAAAGCATCGGCGCT..... 197
   ::::::::::::::::::::
115 AlaAlaArgArgAlaProAlaAlaArgArgGlyHisGlyArgProGlyLe 131
   ::::::::::::::::::::
198 .....GGTGTCTACGCGCGCGCTTCAGGCAAAATGCGCGCGATTT 237
   ::::::::::::::::::::
131 uArgAspLeuHisValArgValHisGlyAlaAlaGlnGlnArgAspAspA 148
   ::::::::::::::::::::
238 CACCGTGGGGA..... 251
   ::::::::::::::::::::
148 rGProProArgArgArgGlyHisGlyProArgHisGlnProProLeuArg 164
   ::::::::::::::::::::
252 .....CGCGTACTTCAGTCAGT..... 269
   ::::::::::::::::::::
165 ArgArgProGlyGlyProGlyAlaAlaValAlaLeuValAlaGlnLeuArgPr 181
   ::::::::::::::::::::
270 .....CGTGAATTCGCGTTGAGCGCACGACGCAAGTTC 300
   ::::::::::::::::::::
```

```
181 oValGlyLeuArgArgValAlaArgAspAlaArgArgArgArgArgArg 198
301 GAGTTTGAAAGCTAGCACTTGAAGCGGCAACTTAAGCGGGAGAGA 350
198 sPProArgProTyrProArgLeuGlySerGlyAlaLeuAlaArgAlaArg 214
351 ACTGGCGCCCAACCTGATCCAGTTT.....GTGACTGGCGC 391
   ::::::::::::::::::::
215 GlyAlaArgAlaGlyAspArgValGlnLeuGlyProGlyAlaAspGlyAs 231
   ::::::::::::::::::::
392 TGGCGACCGCGCTTCAGCAAAATTCCTGCGCGTATGCGGACCGCTT 440
   ::::::::::::::::::::
231 rAlaHisGlyArgValProArgArgGly.....ArgProGlyAlaVal 246
   ::::::::::::::::::::
441 .....CGCCAT...CTTCGTCAATGCGATGACACCAATCCGCT 476
   ::::::::::::::::::::
246 AlaAlaProProArgHisAspGlnArgArgLeuAspProAlaGlnPro 262
   ::::::::::::::::::::
477 GCGTGGCAACCTTACGCTATTCAGAAAGACCGCCGAGATTTCAAAC 526
   ::::::::::::::::::::
263 ArgSerHisProArgGlyLeuProArgAlaProArgArgGlyProArgAr 279
   ::::::::::::::::::::
527 GCGGCGCTGTGGATTCAGCGCTTACCGCAAGCAAAATCCATGCTTGT 576
   ::::::::::::::::::::
279 gArgAspArgGlyValAspLeuValAspArgProProAsp..... 292
   ::::::::::::::::::::
577 AAGGCAGCTGGCGCAGACGTCGCTGAAATGCTGCCAATCATGAAC 626
   ::::::::::::::::::::
293 .....ArgGlyArgArgProGlyVal.....AlaGlnHisPro... 303
   ::::::::::::::::::::
627 ACATGAATTCGCGCGCGCCCATTCGCGCTTGAAGTGCACGACATTC 676
   ::::::::::::::::::::
304 .....LeuArgProAla.....ArgGluPr 311
   ::::::::::::::::::::
677 ATTTCATGACGCGGTGCGCGCAATTAACCGTGTGACATCATTAAT 726
   ::::::::::::::::::::
311 oAlaHisLeuArgAlaArgArgGlyAlaGlyAlaValSerAspArgGlyP 328
   ::::::::::::::::::::
727 CAAAGATGAATTAACATTCGCGCTTGTTCGAACAGCGCGCTGCAAC 776
   ::::::::::::::::::::
328 rArgArgArgProHis.....ArgArgAspArgArg 338
   ::::::::::::::::::::
777 CGAGCGCGATGTCGCGTGAAGTGTTCGACGCAACACCGCGCTGT 826
   ::::::::::::::::::::
339 ArgAlaArgGlyLeuAlaArgArgGlyAlaAspProGlyAlaValProG 355
   ::::::::::::::::::::
827 TGGCTACCGTTTGGTGGCAAGTATCGCAAAATTAATCGCGCGCAATG 876
   ::::::::::::::::::::
355 uAla..... 356
   ::::::::::::::::::::
877 GTTGACACAGACAACCGCGTATTCGCGTTCGATTTGA..... 917
   ::::::::::::::::::::
357 .....ProHisAspArgArgAlaAlaValGlnsAspArgArg 368
   ::::::::::::::::::::
918 .....CGGCGCGATTAC.....ACAAGCGCGCAGC 943
   ::::::::::::::::::::
369 ProGlyProLeuLeuArgArgArgAspHisArgAlaAlaArgAlaHisAr 385
   ::::::::::::::::::::
944 ATTATTTGGAGCGCTACACCAATCAGATTTCGTTATCGAAGA...AGCG 990
   ::::::::::::::::::::
385 gProSerGlyGluAspProArgLeuProHisArgAlaAlaArgArgAspArg 402
   ::::::::::::::::::::
991 CGCAGCAAAAGCTGTCGCGTGGGTTGCGCC..... 1022
   ::::::::::::::::::::
402 lyArgProArgAlaAlaProLeuGlyArgAlaGlyArgGlyGlyGly 418
   ::::::::::::::::::::
1023 .....CGAGCGGCAAAATATCATCATCCAGCGCTA 1051
   ::::::::::::::::::::
419 AspArgSerValGlyArgGlyAlaProGlyArgValAlaArgArgArgArg 435
   ::::::::::::::::::::
1052 CAAC.....CTTCGCGCATTTCTGTAACAAACATCTTC 1086
   ::::::::::::::::::::
435 gArgArgTPrArgArgAlaProArgLeuArgProGlu...GluAlaAlaG 451
```

```

1087 AAGTTCACACAGCCGCTCAACGGCGGCGCCATGTCGCGATTGG 1136
      |||::: ||| |||||::: |||
451 lYvalHisprOarGlyArgProGlyAla..... 465
1137 TACTTACAGCGCGGTATGCGCTTGATATCTGCCACCGCTTTTGC 1186
      |||::: ||| |||::: |||
466 .....AlaGluArgGlnGlnGlyProArgArgProAlaGlySe 480
1187 GCGATTATCTCGCGCATACCGACAGCC..... 1217
      |||::: ||| |||
480 rArgArgGlyArgAlaAlaArgArgGlyArgProAlaAspArgAspA 497
1218 .....CGAGCATTTGGATTGCTTGAATTGGAGCA 1247
      ||| ||||| ||| |||||
497 rGAlaAlaHisArgLeuGlyAlaArgGlyGlyLeuAlaGlyArgGlyArg 513
1248 AGAAG.....CTCGCTTTGTGCAGCTTGCTGCCGCGCAAT 1288
      ||| |||::: |||::: |||
514 ArgArgAspArgGlnProLeuArgAlaArgLeuHisLeuAlaPro.. 529
1289 ACGAATACGGCGCGCTGTGGCAAGTCT 1319
      ||| ||| ||| |||::: |||
530 .ArgAlaArgAlaAlaProAlaArgArgAla 539

```

seq_name: /cgn2.6/prodata/1/aa/5B_COMB.pep:us-09-413-814-78

seq_documentation_block:

```

; Sequence 78, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberger, Steven L
; APPLICANT: Hoefle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; EARLIER FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 78
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-78

```

alignment_scores:

```

Quality: 138.00 Length: 530
Ratio: 0.723 Gaps: 21
Percent Similarity: 36.038 Percent Identity: 23.019

```

alignment_block:

us-09-303-518d-125 x us-09-413-814-78 ..

Align seg 1/1 to: us-09-413-814-78 from: 1 to: 882

```

9 AATCAAAAAAGGTTAAACCTGCGCCATGCGGCGAGACCGAGCAACCG 58
  ::::::::::: ||| |||||::: |||
359 SerArgArgArgArgSerProAlaArgThnProGlyProArgSerProAr 375
59 TTTCAGAGCGCGCGCATACGAAAGTCGGTTCCTGG..... 98

```

```

375 gArgArgProSerProSerAlaArgSerProAlaIleTrrGluArgPro 392
99 .....CGA 101
392 roArgArgProArgAspArgArgProArgArgLeuArgProAlaArg 408
102 AGAATATCGCGGTATGCGCGCTTCGATGAAATGCAAGAGGAGGATCGG 151
409 ArgLeuAlaArgArgAspProGlyAspValProAspProArgAlaIleAr 425
152 TCAAAAAGCGCAAGCGCTTTGAGACAAAGAAATCGCGCGCGTG 201
425 gArgArgAlaProAlaIleArgAlaLeuProGlyAlaAspGlyArgGlyA 442
202 TTACTGCGCGCGCTTCAGGCAAAATCGCGCGATTCACCGTCGCGAAA 251
442 laArgGlyProGly.....ArgArgArgAlaProArgArgGly 454
252 GCGCGTACTTCAGTCAAGTCGTATGCGCTTGAGGCAACGACGAATCG 301
455 AlaAlaAlaArgValProProAspArgAlaAspProProArgArgProAl 471
302 AGTTGAAGCTACGCGACCTGAAGCGCT..... 329
471 aProAlaValAlaArgAlaIleProAlaValAlaArgArgSerAlaArgA 488
330 .....GCGAAA 335
488 laGlyGlnProGlyValGlnHisProValArgGlyAlaArgAspArgPro 504
336 CTTAACGGCGGAGAGTATGCGCGCAACCTGATCCATCGGTTTGTGA 385
505 AlaArgArgArgArgAlaProProGlnProValArgAspArgAlaIleAr 521
386 CTGCGCTGCGCAACCGCTCGCTTCAGCAAAATCTCGCTCATGCGCGAG 435
521 caArgGlyAlaAlaAspAspValGlnArgGlnGly..... 532
436 CGGTTGCCATCTTCGTCATATGGATGACACCAATCGCTGCG..... 479
533 .....ArgArgAlaValProGlyArgArgAlaArgGlyProGly 545
480 .....TGCGACCTACGGTCAATTCGAAGAGC.....CG 511
546 AlaValProAspValArgProArgAlaProArgGlyArgGlyProArgAr 562
512 CCGAGGATTTCAACCGCGCTGTGTGT.....ATTGAGCGGT 549
562 gGlyGlyLeuGlyAlaArgProArgGlyLupProArgAlaValArgProG 579
550 TTGACCGAAGCGCAAAATCCATGTTTGTAAAGCACTGCGCGCAGAGTGGC 599
579 luProArgAlaAlaProPro..... 585
600 GTCTGAATAATGCTGCCAATCATGCAAAACATGATTCGGCGCGCATC 649
585 ..... 585
650 CTGCGGTTTGAAGTGGCAGCACATTCATTTCAGAGCGGTCGCGCG 699
586 .....ArgAlaGlyHisProG 591
700 AATTAACCGGTGTGACCATCAATTATATCAAGATGTAATTACATTGGCCG 749
591 lu..... 591
750 TTTGTTGCACAGCGCGCTGACACCGAGCGCGGTGATGCGCTAGTGTG 799
592 .....AlaProArgArgAlaArgHisArgAlaArgAlaArgAlaProArg 605
800 GTTTCAGTCAACAAACCGCGCGCTTTCGCTACCGTTTGGGTGCGAAA 849

```

```

606 .....A 606
850 GATGCGCAATTAAGCGGGAATGGTGTGACAGCAACCGCGTAT 899
    : : : : :
606 rgvalatrgatgvalatglaag..... 613
900 TTCCGGTTCGATATGACGCGGATTTACAGAGCGCGCGATTTAT 949
    : : : : :
614 ..ArgleuAtgAtgAtgAlaAtgAtgAlaLeuAtgAtgAlaGla 629
950 TGGAGCGCTACCAATCAGATTC.....CGTTATCGAA 984
    : : : : :
629 yProAlaPheProAlaAlaGlaProGlyAlaValAtgAtgLeuAtgA 646
985 GAGCGCGCGCAAGAGCTGTGG..... 1010
    : : : : :
646 rGserProAlaGlyAlaValAlaArgAtgGlyProGlyAlaArgAlaPro 662
1011 .....CTGGGTTCGCGCGACGCGCGCAAAATATCTCA 1042
    : : : : :
663 ValLeuAspArgGluAlaLeuGlyArgAlaAlaProGlyPro..... 676
1043 TCACGCGTACAACTCGCGCATTTCTGAAACAACTCTTCA...G 1089
    : : : : :
677 .ArgAlaGlyArgProAtgAlaGlyAlaLysAspValAlaAtgGlyA 693
1090 TTCACACAGCGCTCAGCGCGCGA..... 1115
    : : : : :
693 laAlaGluProAtgAtgAtgAlaGlyAlaAlaHisProProAspGluAl 709
1116 .....CCGCGCCATGTGCGCGATTTGTTACTTACG 1144
    : : : : :
710 LeuAtgAlaGlyGlyAspLeuProLeuHisGlyProAlaAtgAtgValG 726
1145 AGCGCGTATGCGCTTGGATATCTGCGCGCGCGCTTTTTCG...C 1188
    : : : : :
726 naAlaAtgProAla.....ProAlaHisAtgAlaAtgGlyProAtgA 740
1189 GATTAAATCGTCGCGCA.....TACGCGACGCGCGCGCGCGCTT 1226
    : : : : :
740 rGluHisAtgAtgAtgGluProGluProAtgAtgAspAtgAlaAspAs 756
1227 GGGTTCCTTGAATGACAGAGAGACCTGCTTGTGACGCTTCGCT 1276
    : : : : :
757 ArgLeuLeu.....ArgGluPro.....AlaAtgAl 765
1277 GCCCGCGCAATACGATACGCGCGCTGTGCGCAAGT 1316
    : : : : :
765 aProAspArgLeuAtgAtgArgProAspValAtgAtgAla 778
seq_name: /cgn2_6/ptodata/1/aa/6B_COMB.pep:US-09-105-537-2
seq_documentation_block:
: Sequence 2, Application US/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600,438051
: CURRENT APPLICATION NUMBER: US/09/105,537A
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 5215
: TYPE: PRT
: ORGANISM: Streptomyces venezuelae
US-09-105-537-2

```

```

alignment_scores:
  Quality: 105.50      Length: 601
  Ratio: 0.425        Gaps: 31
  Percent Similarity: 41.265      Percent Identity: 21.238

```

alignment_block:

US-09-303-518D-125 x US-09-105-537-2

Align seg 1/1 to: US-09-105-537-2 from: 1 to: 5215

```

31 CCCATCGCGCGGACGACGCGGACGCGCTTAC..... 63
    : : : : :
1602 ProValPheAlaAlaLeuLeuSprThrValTYrAlaAlaLeuAspArgH 1618
64 ...GACGCGCGCGCATTCAGAGAGTTCGCTTGGCGAAGATAT 108
    : : : : :
1618 sLeuAspArgPro...LeuArgGluLeuValAlaAlaGlyGluGluLeuA 1634
109 .....GCGCGTATCGCGCGCTCGATGAAATCGCAAGAGCGCATGCC 150
    : : : : :
1634 sPheThrAlaTyThrGlnProAlaLeuPheAlaPheGlu..... 1647
151 GTCAAAAAGGCCAAGTCTGTTGACACAAAAGATCCGCGCTGCT 200
    : : : : :
1648 .....ValAlaLeuPheArgLeuLeuGluHisGlyLeuValA 1660
201 G.....TTACTGCGCGCGCTCAGGCAAAATCGCGCGATTCAC 240
    : : : : :
1660 lProAspLeuLeuThrGlyHisSerValGlyGluLeuAlaAlaHisS 1677
241 .....CGT 243
    : : : : :
1677 alAlaGlyAlaLeuSerLeuAspAlaAlaArgLeuValThrAlaArg 1693
244 GCGGAAAAGCGGCTACTCAGTCAGTCGATTCGCTGGAAGGCAACA 293
    : : : : :
1694 Gly.....ArgLeuMetGlnSer.....AlaArgGluGlyAla 1705
294 CGAATCGAGTTCGAGCGCTACGACGCGTGAAGCGTGAAGCTTAAGCG 343
    : : : : :
1705 aMetIleAlaValGlnAlaGlyGluAlaGlyValAlaGluSerLeuLysG 1722
344 GCGAAGAGTGCSC.....CGCAACTG 366
    : : : : :
1722 lTyThrGluArgValAlaValAlaAlaValAsnGlyProThrAlaVal 1738
367 ATCCATTCGCGT.....TTGGGAC 386
    : : : : :
1739 ValValSerGlyAspAlaAspAlaAlaGluGluIleAtgAlaValThrAl 1755
387 T.....GCGCTGCGGACCGCTCGCTTC..... 408
    : : : : :
1755 aGlyArgGlyAlaArgThrArgAtgLeuArgValSerHisAlaPheHis 1772
408 ..... 408
1772 eProHisMetAspAspValLeuAspGluPheLeuArgValAlaGluGly 1788
409 .....AGCAAAATTCCTCGCGGATCGGATCGGCGCTTGC 443
    : : : : :
1789 LeuThrPheGluGluProArgGlieProValValSerThrValThrGlyAl 1805
444 CATCTTCGTCATCGATGAGACACCAATCG..... 474
    : : : : :
1805 aLeuValThrSerGlyGluLeuThrSerProAlaTyThrValAspGlnI 1822
475 .....CTGGCTGCGGAC 486
    : : : : :
1822 leArgArgProValArgPheLeuAspAlaValArgThrLeuAlaAlaGln 1838
487 CTTACGCTATTTATCAAGAGCGCGGAGGATTTCAACGCGCGCTGTT 536
    : : : : :
1839 AspAlaThrValLeuValGluIleGlyProAsp.....Al 1850

```

537 GGTATTGAGCGGTTTGACGACGAAATCCATGTTTGAAGCAGCTG 586
1850 aValLeuThrAlaLeuAlaGluAlaLeu.....AlaProG-1863
587 GCGCAGCGTGGCG...TGTGAAATGCTGCCACATCGAACAACATGAA 633
1863 lYThrAspAlaProAspAlaArgAspValThrValValProLeuLeuArg 1879
634 TTGGCGCGCGCGCATCCT.....GCCGGTTGAGTGGCAGCA 671
1880 AlaGlyArgProGluProGluThrLeuAlaAlaGlyLeuAlaThrAla 1896
672 CATTCAT.....TTTCATGACGCGCGCGCGGA 700
1896 sValHisGlyAlaProLeuAspArgAlaSerPhePheProAspGlyArg 1913
701 ATGAAACCGTGGACCATCAATTCATGATGATTAATTCATTTGCCCT 750
1913 rGThrAspLeuProThrTyraAlaPheArg.....GluHis 1925
751 TTGTTGCAACAGCGCGTGTGAACCGCGCGTATGCGCTAGGTGG 800
1926 TyTrpLeuThrProGluAlaArgThrAsp..AlaArgAlaLeuGlyPh 1941
801 TTCTCAAGTCACAAACCGCGCTTGGTACCGTT...TTGGGTGCGA 847
1941 eAspProAlaArgHisProLeuLeuThrThrValGluValAlaGlyG 1958
848 AAGTATGCAAAATTAATGCGGCGGAAATGGTTGACACGACGACCGCTG 897
1958 lYAspGlyValLeuLeuThrGlyArgLeuSerLeuThrAspGlnProTrp 1974
898 ATTTCCGGTGGTATGTAACCGCGCGATTAACAGCGCGCGACGATTA 947
1975 LeuAlaAspHisMetValAsnGlyValValLeuLeuProAlaThrAlaPh 1991
948 TTGG..... 951
1991 eLeuGluLeuAlaLeuAlaAlaGlyAspHisValGlyAlaValArgValG 2008
952GGACGCTAC 960
2008 lUGluLeuThrLeuGluAlaProLeuValLeuProGluAlaGlyAlaVal 2024
961 CACATTCAGATTCGTTATCGAAGAACGCCGACGAAA.....GAGCT 1004
2025 ArgIleGluValGlyValSerGlyAspGlyGluSerProAlaGlyArgTh 2041
1005 GTTCGGTGGGTGGCGCGCGACGCAAAATCTCATC..... 1044
2041 rPheGly...ValTyrSerThrProAspSerGlyAspThrGlyAspAspA 2057
1045ACGCGTACAAACCTCGGC..... 1062
2057 lAProArgGluTrpThrArgHisValSerGlyValLeuGlyGluGlyAsp 2075
1063CATTTCTGAAAAACAACACTTTCAGTTCAA 1094
2074 ProAlaThrGluSerAspHis..ProGlyThrAspGlyAspGlySerAla 2090
1095 CACAGCGCGTCAACGCGCGCGACCGCATGTCGATGGTGGTATAG 1144
2090 lAtrProProAlaAlaAlaThrAlaThr..ProLeuAspGlyValTyrAs 2106
1145 AGCGCGTATGCCCTTGATATCCCTGCCACCCCTTTTGGCGCATTTA 1194
2106 rArgLeuAlaGluLeuGlyTyrGlyTyrGlyProAlaPheGln..... 2120
1195 ATCTGCGCGATACGACGCGCGGACGATTTGGGTCTGTAATGGA 1244
2121GlyLeuThrGlyLeuTrp 2126

1245 CGA...AGAAGACCTGCTTTGTGACAGTTGCGTGTGCC..... 1280
2127 ArgAspGlyValAspThrLeuAlaGluIleArgLeuProAlaAlaGlnH 2143
1281GGCAATACGATACGCGCGCTGTGGCAACTGCT 1319
2143 sGluSerAlaGlyLeuPheGlyValHisProAlaLeuLeuAspAlaAla 2159
seq_name: /cgn2_6/ptodata/1/laa/6B_COMB pep: US-09-156-836B-2
seq_documentation_block:
; Sequence 2, Application US/09156836B
; Patent No. 6242585
; GENERAL INFORMATION:
; APPLICANT: Sivastava, Ranjana
; APPLICANT: Kumar, Deepak
; APPLICANT: Srivastava, Brahm Shanker
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SPECIFIC DNA FRAGMENT
; FILE REFERENCE: U 011876-4
; CURRENT APPLICATION NUMBER: US/09/156,836B
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 08/997,897
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 430
; TYPE: PRN
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (4)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (6)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (20)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (29)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (54)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (64)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (69)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (89)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (99)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (114)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (119)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (129)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (159)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE
; LOCATION: (169)
; OTHER INFORMATION: amino acid has not been identified
; NAME/KEY: UNSURE

1203 CGATACCGACAGCGC 1217
 |||||
 404 ArgTyrArgHisArg 408

seq_name: /cgn2.6/ptodata/1/aa/6B-COMB.pep:US-09-036-987A-6

seq_documentation_block:

Sequence 6, Application US/09036987A

Patent No. 6143526

GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.

APPLICANT: Broughton, Mary C.

APPLICANT: Crawford, Kathryn P.

APPLICANT: Madduri, Krishnamurthy

APPLICANT: Merlo, Donald J.

APPLICANT: Treadway, Patti J.

APPLICANT: Turner, Jan R.

APPLICANT: Waldron, Clive

TIME OF INVENTION: Biosynthetic Genes for Spinosyn Insecticide

TITLE OF INVENTION: Production

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences LLC Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/036,987A

FILING DATE: 09-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R.

REGISTRATION NUMBER: 28,479

REFERENCE/DOCKET NUMBER: 50,608

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317)337-4816

TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 5588 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-036-987A-6

Alignment_scores:

Quality: 104.50

Ratio: 0.469

Percent Similarity: 41.220

Percent Identity: 20.518

Length: 541

Gaps: 26

Alignment_block:

US-09-303-518D-125 x US-09-036-987A-6

Align seg 1/1 to: US-09-036-987A-6 from: 1 to: 5588

22 CTAAACCTGCC.....ATCGCGGACGACCGGAG..... 51

|||||
 |||||
 2424 LeuSerLeuProAspAlaAlaArgMetValAlaGlyArgAlaArgLeu 2440

52 .CAAGCGGTTACGAGCGCGCGGCATACGAGAGTTCGCTTGGCG 100

|||||
 |||||
 2440 tGlnAlaLeuProSerIleGlyAlaMetLeuAlaValAla...AlaGlyG 2456

101 AAGAAATATGCCGTTATGCCGCTCGATGCAAGTCAAGGAAGCGCATGCC 150

|||||
 |||||
 2456 IuGluGln.....LeuArgProLeuLeu.....AlaAspArg 2466

151 GTCAAAAAGGCCAAGTGCCTGTTGAAAGACAAAAAGAAATCCGGCGTGT 200

|||||
 |||||
 2467 ValAspGlyAlaGlyLe.....AlaAl 2474

201 GTTACTGCGCGCGGCTTACGCAAAATCGCGCATTCACGCGGCGAA 250

|||||
 |||||
 2474 aValAsnAlaProGluSer.....ValValLeuSerGlyAsp 2487

251 AGCGCTACTTACATGATGTCGATTCGCGGTGAAGCAACGAC..... 294

|||||
 |||||
 2487 rGluValLeuAspAspIleAlaGlyAlaLeuAspGlyGlnGlyLeArg 2503

294

2504 TrpArgArgLeuArgValSerHisAlaPheHisSerTyrArgMetAspPr 2520

295

2520 oMetLeuGlnIuPheAlaGluIleAlaArgSerValAspTyrArgArg 2537

314 ACGACCTGAGCGCGTGGCAACTTAAGCGCGAA..... 348

2537 LysPheLeuProValValSerThrLeuThrGlyGlyLeuAspThrAlaGly 2553

349

2554 ValMetAlaThrProGluTyrTrpValArgIleValArgIuProValArg 2570

358

2570 gPheAlaAspGlyAlaArgValAlaLeuAlaGlnGlyValAlaThrIleP 2587

392 TCGCGACCCGTCGTCGTCGCAAAATTCGCGCGCATTCGCGCGCTTC 441

|||||
 |||||
 2587 heGluLeuGlyProAspAlaThrLeuSerAlaLeuIleProAspCysHis 2603

442 GCCATCTTCGTCATGCGATGACACCAATTCGCGCTGCGCGACCTAC 491

|||||
 |||||
 2604 SerTrpAlaAspGlnAlaMetProIleProMetLeuArgLysAspArg 2620

492 G.....GTCATTAATCAAGAACCGCGGAGGATTTAAACGGCGC 532

|||||
 |||||
 2620 rGluThrIuThrValAlaAlaValAlaAlaArgAlaHisThrArgGly 2636

533 TGTGGTATGAGCGGTTGACGCAACGCAAAATTCATGTTGTAAAGCA 582

2636

583 GCTGGCGACAGCGTCGCGTGAATAATGTCGCAACATCGAACAACATGA 632

|||||
 |||||
 2637

633 ATTCGCGCGCGCGCATCTGCGGTTGAGTGGACGACGACATTCATTGA 682

|||||
 |||||
 2644

683 TCGAGCGCGTGGCGCGCAATMAAACGCTGTG.....ACCATCAATTAT 726

|||||
 |||||
 2656 roThrTyrAlaPheGlnArgGlnArgTyrTrpLeuIuThrSerAspTyr 2672

727 CAAGTATTAATTACATGCGCGCTTGTGTCACACGAGCGCTGTGAACAC 776

|||||
 |||||
 2673 GlyAspValThrGlyIleGlyLeuAlaAlaAlaGlu..... 2684

777 CGAGCGGCTGATTGCCCTAGTGCTTCGCAAGTCAACAAACGCGCTCT 826

2684

827 TCGTACCGTTTGGTGGCAAGATATCGCAAAATTAATCGCGCGCAATTG 876

|||||
 |||||


```

seq.name: /cn2_6/ptcddata/1/iaa/6A_COMB.pap:US-09-335-409-5
: seq.documentation block:
: Sequence 5, Application US/09335409
: Patent No. 6121029
: GENERAL INFORMATION:
: APPLICANT: Schnupp, Thomas
: APPLICANT: Ligou, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goerlach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/09/335,409
: CURRENT FILING DATE: 1999-06-17
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 5
: LENGTH: 7257
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
: US-09-335-409-5

```

[illegible]

```

6504 AspCysArgGluLysIleGluProAlaGluLysArgProPheArgLeuGlu 6520
450 CGTCAAT.....GCGATGGACA 466
6520 uilleaspGlySerGlyValLeuAspAspLeuValLeuArgAlaThrGlu 6537
467 CCAATCCGGTGGCTGCCAGCCCTACGGTCATTATCAAGAACGCCGCGAG 516
6537 rGArgProProGluProGluValGluIleAlaValAlaAla... 6552
517 GATTTCAAACGCCGCTGTGTGATGACCGCTTGGACCAAGCAACAAATTC 566
6553 .....GlyLeuAsnPheLeuAspValMetArgAlaMetGlyIle 6565
567 CCAT.....GTTGCTAAG. 579
6565 eTyTProGluProGluAspGlyProValAlaLeuGlyAlaGluCysSerG 6582
580 .....GCAGCTGGCGGAGAGCTGCGCTT.....GAA 606
6582 lYArgIleValAlaMetGlyGluGlyValGluSerLeuArgIleGlyGln 6598
607 AATGCTGCCAATCGAACAACATGATTCGCG..... 639
6599 AspValValAlaValAlaProPheSerPheGlyThrHisValThrIleAs 6615
640 .....GCCCGCATCTCCGCTTGGATGGACGACGACATTC 676
6615 pAlaArgMetLeuAlaProArgProAlaAlaLeuThrAlaAlaGlnAla 6632
677 ATTTCATGACGCGGTCGCGCAATAAACGCTGTGACATCAATTAT 726
6632 laAlaLeu...ProVal...AlaPheMetThrAlaTrp..... 6642
727 CAAAGCTAATTAACATTCGCGCTTGTGTCAAACGCGCTCTGAACAC 776
6643 TyTcIlyLeuValHisLeuGlyArgLeuArgAla.....G1 6654
777 CGACCGCGTAT.....GCCCTAGTGGTCTCAAGTCACAAAC 817
6654 yGluArgValLeuIleHisSerAlaThrGlyGlyThrGlyLeuAlaIle 6671
818 CGCGCTCTTCCGTCACGCTTTGGTGGCAAGTA..... 852
6671 alGlnIleAlaArgHis...LeuGlyAlaGluIlePheAlaThrAlaGly 6686
852 ..... 852
6687 ThrProGluLysArgAlaTrpLeuArgGluGlnGlyIleAlaHisValMe 6703
853 .....TCGCAAT.....ACTGCG 868
6703 tAspSerArgSerLeuAspPheAlaGluGlnValLeuAlaIleThrLys 6720
869 GCGAATGTGTGACACAGACAAACGCGTGAATTCGGTGTGATGTAAC 918
6720 yGluGlyValAspValValLeuAsnSerLeuSerGlyAlaAlaIleasp 6736
919 GCGCGGATTACAAAGCGCGCAGATTTTGGAGCGTACACAAATCA 968
6737 AlaSerLeuSer..... 6740
969 GATTTCGCTATCGAAGAGCGCGCAGCAAGACGCTTGGCGTGGGTTG 1018
6741 .....ThreValAlaProaspGlyArgPheIleGluLeu..... 6751
1019 CGCGCGAGCGGACAAATCTCATCAGCGCTACAAACCTCGCGCATTC 1068
6752 ...GlyLysThrAspIleTyTAlaAspArgSerLeuGlyIleAlaHisPhe 6767
1069 CTGAAACAAACCTCTCAAGTTCACACAGCGCTCAAGCGCGCGCGC 1118

```

```

6768 ArgLysSerLeuSerTyT.....SerAlaVal.....AspLe 6778
1119 CGCCATGTGTCGCGATGTACTTACGAGCGCGT..... 1152
6778 uAlaGlyLeuAlaValAlaArgArgProGluArgValAlaAlaIleLeuAla 6795
1153 .....ATGCCCTTGAT 1164
6795 lValValAlaAspLeuLeuAlaArgGlyAlaLeuGlnProLeuProValGlu 6811
1165 ATCTGCGCCACCGCTCTTTCGCGATTTAATCGTCGGCGATACCGACAG 1214
6812 lIlePheProLeuSerArgAlaAlaAspAlaPhe...ArgLysMetAlaGln 6827
1215 CGCGAGCATGTGGTGTCTTGAATTGGACGAAGAACCTCGCTTGT 1264
6828 AlaGlnHisLeuGlyLysLeuValLeuAlaLeuGlnLysProAsp...Va 6843
6843 lArgIleArgValProGlyGln 6850

```

seq_name: /cgn2_6/ptodata/1/aa/5b_COMB.pep:US-09-568-102-5

seq_documentation_block:

```

; Sequence 5, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; PRIOR FILING DATE: 2000-05-10
; CURRENT APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: prt
; ORGANISM: Sorangium cellulosum
; US-09-568-102-5

```

alignment_scores: Quality: 103.50 Length: 541
Ratio: 0.452 Gaps: 27
Percent Similarity: 42.329 Percent Identity: 21.442

alignment_block:

US-09-303-518D-125 x US-09-568-102-5 ..

Align seg 1/1 to: US-09-568-102-5 from: 1 to: 7257

```

55 GCGCTTACAGAGCGCGCCATTACGAAAGTCGCTTGGCGGAA... 102
||| ..... ||| ..... ||| ..... ||| .....
6354 AlaSerPheAspGlyGlnAlaProThrSerValValHisLeuIleYerie 6370
103 GAATATGCGGTATGCGCCCTCGATGAAGTCAAGGAAGCGATCGG 151
6370 uAspGluArgGlyValLeuAspAlaAspAlaProPheAspAlaAla 6387
152 TCAAAAAAGC.....CAAGTCTGTGTAAGACAA 183
6387 euGluGluSerLeuValArgGlyCysAspSerValLeuTrpThrValGln 6403
184 AAGAATCGCGCGCTGTCTTACTGCGCG.....GCTTC 218
||| ..... ||| .....

```

```

6404 AlaValAlaGlyAlaGlyPheArgAspProProArgLeuThrIleValTh 6420
219 AGGCAAAATCGCGGATTCACCTGGCGAAACCGGTACTTCAGTACG 268
6420 rArgGlyAlaGlnAlaIleGlyAlaGlyAspValSerValAlaGlnAla 6437
269 TC.....GTGATTCCGTTGAA..... 285
6437 rOLEuLeuGlyLeuGlyArgValIleAlaLeuGlnIleAlaGlnLeuArg 6453
286 .....GGCAACGACGAAATCGAGTTTGACG 311
6454 CysAlaArgIleAspLeuAspProAlaArgArgAspGlyAlaValAspG 6470
312 CTACGCACTGAACGCTGGCAACTTAAGCGCGAAGAGTCCGCCGCA 361
6470 uLeuLeuAlaGlnLeuLeuAlaAspAlaGlnGlnGlnValAlaIlePhe 6487
362 ACCTGATCCAAATCGGTTTGATGACGCTGGCACCCTCCGTTGACG 411
6487 rGgGlyGlyAlaArgValAlaArgLeuValAlaArgArgLeuProGlnThr 6503
412 .....AAATCTCGGCTCGATGGCGACCGCTTGCCATCTT 449
6504 AspCysArgGlyLysIleGlnProAlaGlnGlyArgProPheArgLeu 6520
450 CGTCAT.....GGATGGACA 466
6520 uIleAspGlySerGlyValLeuAspLeuValLeuArgAlaThrGln 6537
467 CCAATCGCGCTGGCGGACCTACGCTATTCATCAAGAACCGCCGAG 516
6537 rGArgProProGlyProGlyGlnValGlnIleAlaValGlnAla 6552
517 GATTTCAAACGCGCTGTTGATGAGCGCTTGACGCAACGCAAAAT 566
6553 .....GlyLeuAsnPheLeuAspValMetArgAlaMetGlyIle 6565
567 CCAT.....GTTGTAG 579
6565 eIyTrProGlyProGlyAspGlyProValAlaLeuGlyAlaGlnCysSer 6582
580 .....GCAGCTGGCGACAGCGTCCGCT.....GAA 606
6582 yArgGlyLeuAlaMetGlyGlnGlyValGlnSerLeuArgIleGlyGln 6598
607 AATGCGCAACATCGAAACACATGAATTCGCG..... 639
6599 AspValAlaValAlaProPheSerPheGlyThrHisValThrIleAs 6615
640 .....GGCCGCAATCGCGGTTTGAGTGGCAGCAGCATTC 676
6615 pAlaArgMetLeuAlaProAlaArgProAlaIleLeuThrAlaGlnIle 6632
677 ATTTCATCGACCGGCTGGCGGATTAACCGTGGACCATCATATTA 726
6632 lAlaIle...ProVal...AlaPheMetThrAlaTrp..... 6642
727 CAAGATGAATTAACCATGGCGCTTTGTTTCAACAGCGCGCTGAAC 776
6643 TyrGlyLeuValHisLeuGlyArgLeuArgAla.....G 6654
777 CGAGCGCTGAT.....GCCCTAGGTGTTCTCAAGTCAACAAC 817
6654 yGlnArgValLeuIleHisSerAlaThrGlnGlyThrGlyLeuAlaVal 6671
818 CGCGCGCTGCTGACGCTTTGGGTCGGAAGTA..... 852
6671 aGlnIleAlaArgHis...LeuGlyAlaGlnIlePheAlaThrAlaGly 6686
852 ..... 852
6687 ThrProGlyLysArgAlaTrpLeuArgGlnGlyIleAlaHisValMe 6703

```

```

853 .....TCGCAAAAT.....ACTGCG 868
6703 tAspSerArgSerLeuAspPheAlaGlnGlnValLeuAlaIleThrIle 6720
869 GCGAATTTGGTTGACACAGACACACCGGCTGATTTCCGTTCCGATTAAC 918
6720 yGlnGlyValAspValValLeuAsnSerLeuSerGlyAlaAlaIleAsp 6736
919 GCGCGATTCACACAGCGCGCACGATTTTGGAGCTTACCACAAATCA 968
6737 tAlaSerLeuSer..... 6740
969 GATTTCGTTATCGAAGAGCGCGCAAGAGCTGTCGCTGGGTTG 1018
6741 .....ThrLeuValProAspGlyArgPheIleGlnLeu..... 6751
1019 CGCGCGACCGCGACAAATTAATCCATCAGCGGTACAACTCGCCATTT 1068
6752 ..GlyLysThrAspIleTyrAlaAspArgSerLeuGlyLeuAlaHisPhe 6767
1069 CTGAACAACTCTTCAAGTTCAACACAGCGCGTACACGCGGACCG 1118
6768 ArgLysSerLeuSerTyr.....SerAlaVal.....AspLe 6778
1119 CGCCATGTCGCGATTTGATTCAGACGCGCTG..... 1152
6778 uAlaGlyLeuAlaValAlaArgProGlnArgValAlaIleLeuAlaG 6795
1153 .....ATGCCCTTGAT 1164
6795 lValValAspLeuLeuAlaArgGlyAlaLeuGlnProLeuProValGln 6811
1165 ATCTTCCGACCCCTGCTTTGGCGATTTAAATCGCGGATTAACGACAG 1214
6812 lIlePheProLeuSerArgAlaIleAspAlaPhe..ArgLysMetAlaGln 6827
1215 CGCGCAGGATTTGGTGGTGGATTTGGACCAAGACCTCGCTTGT 1264
6828 lAlaGlnHisLeuGlyLysLeuValIleuAlaLeuGlnAspProAsp...Va 6843
1265 GCAGCTTCGCTGCGCGGCA 1286
6843 lArgIleArgValProGlyGln 6850
seq_name: /cgn2_6/ptodata/1/iaa/0b_COMB pep:US-09-567-969-5
seq_documentation_block:
/ Sequence 5, Application US/09567969
/ Patent No. 6355457
/ GENERAL INFORMATION:
/ APPLICANT: Schupp, Thomas
/ APPLICANT: Ligon, James
/ APPLICANT: Molnar, Istvan
/ APPLICANT: Zirkle, Ross
/ APPLICANT: Cyr, Devon
/ APPLICANT: Goerlach, Joern
/ TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
/ FILE REFERENCE: 4-30582A
/ CURRENT APPLICATION NUMBER: US/09/567, 969
/ PRIOR APPLICATION NUMBER: 09/335, 409
/ PRIOR FILING DATE: 1999-06-17
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 7257
/ TYPE: PRT
/ ORGANISM: Sorangium cellulosum
/ US-09-567-969-5

```

alignment_scores:


```

1  GENERAL INFORMATION:
2  APPLICANT: Schupp, Thomas
3  APPLICANT: Ligon, James
4  APPLICANT: Molnar, Istvan
5  APPLICANT: Zinke, Ross
6  APPLICANT: Cyr, Devon
7  APPLICANT: Goerlach, Joern
8  TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
9  FILE REFERENCE: 4-30582A
10 CURRENT APPLICATION NUMBER: US/09/568,480
11 CURRENT FILING DATE: 2000-05-10
12 PRIOR APPLICATION NUMBER: 09/335,409
13 PRIOR FILING DATE: 1999-06-17
14 NUMBER OF SEQ ID NOS: 30
15 SOFTWARE: PatentIn Ver. 2.0
16 SEQ ID NO 5
17 LENGTH: 7257
18 TYPE: PRT
19 ORGANISM: Sorangium cellulosum
20 US-09-568-480-5

```

```

alignment_scores:
  Quality: 103.50      Length: 541
  Ratio: 0.442        Gaps: 21
  Percent Similarity: 42.329      Percent Identity: 21.444
alignment block:
US-09-303-518D-125 x US-09-568-480-5 ..
Align seg 1/1 to: US-09-568-480-5 from: 1 to: 7257

```

55 GCGGTTTACGACGGCCCGGCATTACCGAAGTCGGCTTGGCGCA.. 102
 |||:||||| ||| |||:|||||
 6354 Alaserthepsglytinalaprotthrservalvalhstleuglysele 6370T
 103 .GAATATCCGGTATGCGCCCGCCATGATCAAACTCAAGGAGCGATGGCG 151
 |||:||||| :||| :|||:|||||
 6370 uaspglytarglyvalleuaspalaspalaspalaspalaspalaspal 6387T
 152 TCAAAAAGGC.....CAAGTCGTGTTGAAGACAA 183
 :|||:||||| :|||:|||||
 6387 euglucjuserleuvalarglycysaspservaltheutrlthValgln 6403
 184 AAGAAATCCGGCGCTGTTTCTGTCGGCG.....GCTTC 218
 |||:||||| ||| |||
 6404 AlavalalaglyalaclypheargasprproalargleutrlleuValth 6420
 219 AGCCAAATTCGCGCGCGATTCACCGTGGCGAAAGCGCTACTGATGAC 268
 ||| ||| ||| |||: |||
 6420 rargglyalaglnalaleglyalaglyaspvalserlalaclylnlar 6437
 269 TC.....GTGATTCGCGTTGA.. 285
 6437 roleuleuclyleuglyargvalillealaleuglnhslaglyleuValarg 6453
 286GGCAACAGCAATTCAGTTTGACG 311
 :|||:||||| :|||:|||||
 6454 CysAlaargIleaspleuaspvalargargasprglytinalaspgl 6470
 312 CTACGCACTGACGCGCTGGCAACTTAAGCGGCGAAGAGTGGCGCGCA 361
 ||| ||| |||:||||| :||| |||:|||||
 6470 uleuleualaglyleuleualaspalaglyluclylvalalpheh 6487
 362 ACCGTGTCAAATCCGGTTGTGCACTGCGCTGGCGCACCCCTCCGTTACG 411
 :|||:||||| :|||:||||| :|||:|||||
 6487 rgglyglytarglyalalaargleuvalargargleuproglythr 6503
 412AAATTCCTGCGCTGACGCGCGCGCTTGGCCATCTT 449
 ||||| :|||:||||| :|||:|||||
 6504 AspCysargglytlylleglproalaglylYargtrropheargyleu 6520
 450 CGTCAAT.....GGCATGGACA 466

```

6520 uilleaspdysergllyalleuspsarpleuValleatrgalatlnhlna 6537
      ::::: 111 ::::
6521 CCAATCCGCTGGCTGGCCGACGCGTCACTTGTCAAAAGACGGCCGAG 516
      ::::|::| ::::|::| ::::|::| ::::|::|::|::|::|
6537 rgatrgproptrogllyprogllygluvalclutlleVala1clutla1a... 65523
517 GATTTCAAACGCGGCGCTGTGTGTATGAGCGCGTTTGACCGAACGCAAAAT 566
      ||||| 111::|::|::|::|::|::|::|::|::|::|::|
6553 .....gilyeuanphleuApsarValmetAtrgAla1metclY1l 65655
567 CCAT .....GTTTGAAG. 579
      |::|::|::|::|::|::|::|::|::|::|::|::|::|
6565 enytrprogllyprogllyaspldyprovalAla1eugllyalag1ucysserg 65828
580 .....GCAGCTGGCGGACAGCGTCCGCT.....GAA 606
6582 lyatrglleValAla1metclYglu1yvalclutserleuAtrglleclYln 65988
607 AATCGTGGCAACATCGAAACACATGAATTCGGC.....639
6599 AsprvalAla1valAla1arpropheserPhaglythlnh1nsvalth1nleas 6615
640 .....GGCCCGATCCCTCCGCGTGTGAGTGGACGCGACATTC 676
6615 palAtrgMetleuAla1trgtrproAla1leuThra1la1aglnlaa 6632
677 ATTTCATCGACGCGATCGCGCGAAATAAACCGTGGACCATCAATAT 726
6632 la1aleu...Proval...Ala1phemetTh1alatr.....6642
727 CAAGATGTATGTACATTTGGCGCTTTGTTTGCACACGCGCGTGAACAC 776
6643 tyrglyleuvalh1sleu1yAtrg1eulAtrgAla.....gl 6654
777 CGAGCGCGTGAAT.....GCCCTAGGCGGTTCCTCAATCAACAAAC 817
6654 ygluAtrgvalleu1leu1serAla1thelYclYthelYleuAla1ay 6671
818 CGGCGCTTCGCGTACCGCTTTGGTGGCGAAAGA.....852
6671 alclnleuAla1trg1ns...leuclYAla1aglu1lePhelAla1thralagly 6686
852 .....852
6687 Th1ProglulysAtrgAla1trpleuAtrg1us1nlylleAla1n1syalme 6703
853 .....TCGCAAAAT.....ACTGGCG 868
6703 tAsperAtrgserleuApsarPhelAglu1nvalleuAla1Ala1th1lyls 6720
869 GCGAAATGTGTACACACACACCGCGCGATTTCCGGTTCGGTATTTGAAC 918
      ||||| 111::|::|::|::|::|::|::|::|::|::|::|
6720 yglu1yvalAsprvalAla1leu1ns1erleuSerclYAla1Ala1leasr 6736
919 GCGCGCATTCACACAGCGCGCACAGATTAATTGGACGCGTACACAAATGA 968
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|
6737 AlaserleuSer.....6740
969 GATTTCGCTTATCGAAGAAGCGCGCAAAAGACTTTCGGCTGGGTGG 1018
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|
6741 ...Th1leuvalProAsprglYatrgPhel1eglu1eu.....6751
1019 CGCGCGACGCGGACAAATATCTCCATCAACGCGGTTCACAAACCTCGGCGCATTTG 1068
6752 ...clYlyuTh1AsprleuThra1la1AsprAtrgserleuclYleuAla1ns1phe 6767
1069 CTGAAGAAACAAACCTCTCAAGTTCAACACACGCGCTCAACAGGCGGCGACCG 1118
6768 Atrg1ys1erleuSer1ytr.....SerAla1val.....Asple 6778
1119 CGCCATGTGTCGCGATTCAGTACTTAACAGCGCGCTG.....1152
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|

```



```

869 GCGAATTGGTTGACACAGACACCGCGTGTATTCGGTTCGTTGATTCAGC 918
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6720 LyluglyValAspValValLeuAsnSerLeuSerGlyAlaLeuLeasp 6736
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
919 GGGCGCATTAACACAGCGCGACGATTAATTTGGAGCGTACACACATCA 968
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
6737 AlaSerLeuSer..... 6740
969 GATTTCGTTATGAGAGCGCGACGAGAGAGCTGTTCGGTGGTGGT 1018
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
6741 ...ThreValAlProAspGlyArgPheLeuLeu..... 6751
1019 CGCGCGACCGCGACAAATCTCCATCAGCGCTACAGCCCTCGCCCATTC 1068
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6752 ..GlySerThrAspLeuAlaAspArgSerLeuGlyLeuAlaHisPhe 6767
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1069 CTGAACAAACAACTCTTCAGTTCAACACACCGCTCAAGCGCGACCG 1118
      ||||| ::::: ::::: ::::: ::::: ::::: ::::: :::::
6768 ArgLysSerLeuSerTyr.....SerAlaVal.....AspLe 6778
1119 CGCCATGTGGCGATTTGGTACTTACGAGCGCGT..... 1152
      ||||| ::::: ||||| ||||| ||||| ||||| ||||| |||||
6778 ValAlaGlyLeuAlaValArgArgProGluArgValAlaAlaLeuAla 6795
      ||||| ::::: ||||| ||||| ||||| ||||| ||||| |||||
1153 .....ATGCCCTTGGAT 1164
6795 LuValValAspLeuLeuAlaArgGlyAlaLeuGlnProLeuProValGlu 1164
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1165 ATCTGCGCCACCTGCTTTTGGCGATTTAATGTCGCGCATTCAGCAGAG 1214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6812 IlePheProLeuSerArgAlaAlaAspAlaPhe..ArgLysMetAlaGln 6827
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1215 CGCGCAGGACATGGTGGTCTTGCATTTGACGAGAGAGACCTCGCTTGT 1264
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6828 AlaAlaHisLeuGlyLysLeuValLeuAlaLeuGluAspProasp...Va 6843
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
1265 GCACCTTCCTGTCGCGCGCGCA 1286
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6843 ArgIleArgValProGlyGlu 6850

```

seq_name: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:us-09-568-472-5

```

seq_documentation_block:
; Sequence 5, Application US/09568472
; Patent No. 6358719
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zikkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,472
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-568-472-5

```

Alignment_scores:

Quality:	103.50	Length:	541
Ratio:	0.452	Gaps:	27
Percent Similarity:	42.329	Percent Identity:	21.442

alignment_block:

US-09-303-518d-125 x US-09-568-472-5 ..

Align seg 1/1 to: US-09-568-472-5 from: 1 to: 7257

```

55 GCCGTTTACAGCGCGCGCATTTACGAGTGGTTCGTTGGCGAA.. 102
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
6354 AlaSerPheAspGlyGlnAlaProHisSerValAlaHisLeuGlySerLe 6370
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
103 GAATATCGCGGTATGCCCGCCCTCGATGAAAGTCAGAGAGCGCATCCG 151
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6370 uAspGluArgGlyValLeuAspAlaAspAlaProPheAspAlaSplal 6387
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
152 TCMAAAGAGC.....CAAGTCGTTTGAAGACAA 183
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6387 euGluGluSerLeuValArgGlyCysAspSerValLeuTyrPheValGln 6403
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 AAGAATCCGGCGCTGTGTTTACTGCGCG.....GTTTC 218
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
6404 AlaValAlaGlyAlaGlyPheArgAspProProArgLeuTyrPheValTh 6420
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 AGCCAAATCCCGCGCATTCACCGTGGCGAAAGCGGTACTTCACTCAG 268
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6420 rArgGlyAlaGlnAlaIleGlyAlaGlyAspValSerValAlaGlnAla 6437
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 TC.....GTGATTCGCGTGA..... 285
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6437 rGluLeuGlyLeuGlyArgValIleAlaLeuGlnHisAlaGluLeuArg 6453
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 .....GCCACGACGAAATGAGTTGACG 311
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6454 CysAlaArgIleAspLeuAspProAlaArgArgAspGlyValAlaAspI 6470
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 CTACGACCTGAAGCGCTGCAACTTAAGCGCGCAAGAGCGCCCGCA 361
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6470 uLeuLeuAlaGluLeuLeuAlaAspAspAlaGluGluValAlaAlaPhe 6487
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
362 ACCGTATCCATCCGCTTGTGACTGCGCTCGCAGCCGCTGCTCAGC 411
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6487 rGluGlyGluArgValAlaArgLeuValAlaArgLeuProGluThr 6503
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 .....AAATTCCTGCGCGATGCCGAGCGCTTGCATCTT 449
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6504 AspCysArgGluLysIleGluProAlaGluGlyArgProPheArgLeuG1 6520
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 CGTCAT.....CGATGACCA 466
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6520 uIleAspGlySerGlyValLeuAspAspLeuValLeuArgAlaThrGlu 6537
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 CCAATCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 516
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6537 rArgProProGlyProGlyGluValGluIleAlaValAlaAla... 6552
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
517 GATTTCAAACGCGCGCTGTGATTTGAGCGCTTGAACGCAACAAAT 566
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6553 .....GlyLeuAsnPheLeuAspValIleThrGlyAlaMetGlyI 6565
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
567 CCAT.....GTTTGAAG. 579
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6565 eTyrProGlyProGlyAspGlyProValAlaLeuGlyAlaGluCysSerG 6582
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
580 .....GCAAGTCGCGCGAGCGTGGCT.....GAA 606
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6582 LArgGlyLeuAlaIleMetGlyGluGlyAlaGluSerLeuAlaGlyIleGln 6598
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 AATGCGCAACATCGAACAACATGAATTCGCG..... 639
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6599 AspValValAlaValAlaProPheSerPheGlyThrHisValThrLeas 6615
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
640 .....GCCCGCATCCGCGCGGTGATGAGCGCGCGCGCGCATTC 676
      ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
6615 rAlaArgMetLeuAlaProArgProAlaAlaLeuThrAlaAlaGlnAla 6632

```


APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship

441 hrgInvalGlyIleThrLeuLeuProGlnAlaAlaProAlaProIleGln 450

869CCGCAGTATTGCATCTTCGACC 841
||||:|||| ::::~::~:


```

GENERAL INFORMATION:
APPLICANT: Hardy, Daniel M.
TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/276,967
FILING DATE: Submitted Herewith
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSD:418\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2476 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-276-967-2

alignment_scores:
Quality: 99.00 Length: 373
Ratio: 0.615 Gaps: 19
Percent Similarity: 43.164 Percent Identity: 20.912

alignment_block:
US-09-303-518D-125 x US-08-276-967-2 ..
Align seg 1/1 to: US-08-276-967-2 from: 1 to: 2476
381 GTGACGTGGCGTCGCGAC..... 398
1849 LeuAspCysSerAlaHisSerValThrSerCysValProSerCysLeu 1865
399 ..CCGTCGCGTCAG.....CAAAATTCCTGCGTCGATCGCGACG... 437
1865 uProSerCysGlnAspProGlnGlyGlnCysThrGlyAlaGlyAlaPro 1882
438 .....GTTCGCGCATCTT 449
1882 eThrCysGlnGlyGlyCysIleCysGlnProGlyTyrValLeuSerGlu 1898
450 CGTCATGGAGGAGACCAATCCGTCGTCGCGAC..... 485
1899 GlnGlnCysValAlaArgSerGlnCysGlyCysArgAspAlaArgGlyTh 1915
486 .....CCCTACGTCATTAACAAGAGCCGCGAGGATTTCGA.... 524
1915 rPheLeuProValGlyArgPheArgLeuSerSerGlyCysSerGlnMetC 1932
525 .....ACGGCGCGCTTGATTCGACCGTTGACCGAAGCGCA 563
1932 ySValCysThrAlaGlyAlaIleGlnCysArgProPheThrCysProSe 1948

```

```

564 AATCATGTTTGTAA.....GCAGCTGGCGGACGTCGCTGAAA 607
1948 rGlySerGlnCysGlnProAsnGlnAspGlyLysAspPheCysGlnPro 1965
608 ATGTGTCACATCGAACAACATGATTCGGCGCCGATTCCTGCCGGT 657
1965 snSerSerAsnLeu...CysSerValPheGlyAspProHisTyrArgThr 1980
658 TTGAGTGGCGGACGACATTCATTC..... 681
1981 PheAspGlyLeuSerTyrArgPheGlnGlyArgMetThrTyrThrLeu 1997
682 .....ATCGACCGGTC.....G 694
1997 llyThrLeuAspValLeuProAspGlyValGlnProLeuValValGlu 2014
695 GCGCGAATAA.....ACCGTGGACCATCATTAACAGAGTA 735
2014 lYArgAsnLysValTyrProSerLeuThrProValPheLeuGlnGlu 2030
736 ATTACCATTCGGCGCTTGTTCACACAGCGCGTGAACCGAGCGCGT 785
2031 lIeValMetValTyrGlyTyr...ThrValGlnLeuGlnAlaGluLeu 2046
786 GATTGCCCTAGGTGTTCTCAAGTCAACAAACG.....C 820
2046 uLeuValValAsnGlyGlnLysValSerIleProTyrLysProAsnGlu 2063
821 GCCTCTTGCGTACCGCTTTGGTGGCGAAGTATCCAAATTTACTCGGCG 870
2063 yLeuGlnValThrLeuArgGlyArgArgLeuTyrLeuValThrAspPhe 2079
871 GAATTGGTTGACACAGAACCGCGTATTCGGTTCGTTATTCAGCG 920
2080 GluLeu.....ValValSerPheAsnGlyArgAsnAs 2090
921 CGCGATTACACAAGCGCGCAGCATTTATTTGGAGCGCTACCAATACAGA 970
2090 nAlaValIleAlaMetProSerThrTyrLeuGly..... 2101
971 TTTCGTTATCAAGAAGCGCGCAGCAAGAGCTGTTCGCTGGTTCG 1020
2101 ..... 2101
1021 CCGAGCGCGAACAATTCATCCATCAGCGCTACCAACCTGGCGCATTCCT 1070
2102 .....LeuValArgGlyLeuCysGlyAsnTyrAs 2111
1071 GAAATCAAACTCTTCAAGTTCAACACAGCGCGTCAACGCGCGGACGCG 1120
2111 pLysAsnLysArgAsnAspPheMetLeuProAsnGlySerPheThr 2127
1121 CCAT.....GCTGCCGATTGTTACTTACAGCGCGCGATGCC 1158
2127 nasnLeuValPheGlyAsnSerTyrGluValLysAlaLysGlu... 2142
1159 TTGGATATCCGCGCCACCGCTTTTGGCGGATTAATCTCGCGCATAC 1208
2143 .....GlyHisProArgPheSerArgAlaIleArgGlnGlu 2155
1209 CGACAGCGCGCAGGCGATTGGCTCTTGAATTCGAGCAGCAAGACCTCG 1258
2156 GluLysAsnGlnGluSerGlyPheGlnAsnValSerGlyCysSerPro 2172
1259 CTTTGTGCAGCTT 1271
2172 uGlnLeuGlnLeu 2176
seq_name: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:US-08-997-897-2
seq_documentation_block:
; Sequence 2, Application US/08997897C

```

```

1 Patent No. 6114514
2 GENERAL INFORMATION:
3 APPLICANT: SRIVASTAVA, RANJANA
4 APPLICANT: KUMAR, DEEPAK
5 APPLICANT: SRIVASTAVA, BRAHM SHANKER
6 TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SPECIFIC DNA FRAGMENT
7 FILE REFERENCE: u011469-7
8 CURRENT APPLICATION NUMBER: US/08/997,897C
9 NUMBER OF SEQ ID NOS: 7
10 SOFTWARE: PatentIn Ver. 2.0
11 SEQ ID NO 2
12 LENGTH: 430
13 TYPE: PRF
14 ORGANISM: Mycobacterium tuberculosis
15 FEATURE:
16 NAME/KEY: UNSURE
17 LOCATION: (4)
18 FEATURE:
19 NAME/KEY: UNSURE
20 LOCATION: (6)
21 FEATURE:
22 NAME/KEY: UNSURE
23 LOCATION: (20)
24 FEATURE:
25 NAME/KEY: UNSURE
26 LOCATION: (29)
27 FEATURE:
28 NAME/KEY: UNSURE
29 LOCATION: (54)
30 FEATURE:
31 NAME/KEY: UNSURE
32 LOCATION: (64)
33 FEATURE:
34 NAME/KEY: UNSURE
35 LOCATION: (69)
36 FEATURE:
37 NAME/KEY: UNSURE
38 LOCATION: (89)
39 FEATURE:
40 NAME/KEY: UNSURE
41 LOCATION: (99)
42 FEATURE:
43 NAME/KEY: UNSURE
44 LOCATION: (114)
45 FEATURE:
46 NAME/KEY: UNSURE
47 LOCATION: (119)
48 FEATURE:
49 NAME/KEY: UNSURE
50 LOCATION: (129)
51 FEATURE:
52 NAME/KEY: UNSURE
53 LOCATION: (159)
54 FEATURE:
55 NAME/KEY: UNSURE
56 LOCATION: (169)
57 FEATURE:
58 NAME/KEY: UNSURE
59 LOCATION: (182)
60 FEATURE:
61 NAME/KEY: UNSURE
62 LOCATION: (185)
63 FEATURE:
64 NAME/KEY: UNSURE
65 LOCATION: (219)
66 FEATURE:
67 NAME/KEY: UNSURE
68 LOCATION: (259)
69 FEATURE:
70 NAME/KEY: UNSURE
71 LOCATION: (269)
72 FEATURE:

```

1	NAME/KEY: UNSURE
2	LOCATION: (291)
3	FEATURE:
4	NAME/KEY: UNSURE
5	LOCATION: (323)
6	FEATURE:
7	NAME/KEY: UNSURE
8	LOCATION: (339)
9	FEATURE:
10	NAME/KEY: UNSURE
11	LOCATION: (349)
12	FEATURE:
13	NAME/KEY: UNSURE
14	LOCATION: (356)
15	FEATURE:
16	NAME/KEY: UNSURE
17	LOCATION: (366)
18	FEATURE:
19	NAME/KEY: UNSURE
20	LOCATION: (410)
21	FEATURE:
22	NAME/KEY: UNSURE
23	LOCATION: (421)

```
alignment_scores:
  Quality: 98.00
  Ratio: 0.838
  Percent Similarity: 36.909
  Length: 317
  Caps: 16
  Percent Identity: 22.082
```

```
alignment_block:
```

US-09-303-518D-125 x US-08-997-897-2

Align seg 1/1 to: US-08-997-897-2 from: 1 to: 430

[illegible]

```

281 GlycysArgGlyCysArgCysGlnGlyCys***SerHisArgArgPheAr 297
482 ..... 482
297 gCysArgGlyCysCysAsnCysArgCysTrpArgCysArgGlyCysSerA 314
483 .....CGACCTGACGTCATTATCA 503
314 rGArgProGlyLeuProGlyArgAsp***ArgProValGlyHisArgGlyS 330
504 .....AGAAAGCCCGAGAGATTCAACCGGCGCT.. 533
331 IleProThrCysCysPheArgCys***ArgSerProArgSerArgProAl 347
534 GTTGGTATTGACCGCTTGACCGAAGCAACCAAAATCCATGTTGTAAAGCA 582
347 aLeu***TrpProProGlySerCys***ThrAsnProIle..... 360
583 GCTGGCGAGACGTGCCGTCTGAAATGCTGCCAACATCGAAACACATGA 632
361 .....ArgCysCysProSer***Ser..... 367
633 ATTCGGCGCCGCATCTGCCG.....TTGA 661
368 .....ArgProIleProAlaArgProAlaArgLeuProGlyArgSerTyrAr 382
662 GTGG.....CACGCACATTCATTTCATCGA...G' 687
382 gTrpProProThrLysSerGlyArgSerGlnAsnCysTrpHisArgSerS 399
688 CCGGTGCGCGCGAATAAACCGGTGTGACCATCAATTATCAAGA..... 731
399 eArgLysArg.....ThrArgTyrArgHisArgCys***ArgGlyLeuPro 413
732 TGTAAATACATTTGGCGCTTGTGTTGCAACAGCGCTGCAACACCGAGC 781
414 ThrThrAlaHisTrpSerSer***ArgSerArgAlaProGlyAlaAsnAla 430
782 G 782
430 A 430
seq_name: /cgn2_6/protodata/1/1aa/6B.COMB.pep:us-09-370-368-5
seq_documentation_block:
: Sequence 5, Application US/09370368
: Patent No. 6258932
: GENERAL INFORMATION:
: APPLICANT: Anders Vahline
: TITLE OF INVENTION: PEPTIDES THAT BLOCK VIRAL INFECTIVITY
: TITLE OF INVENTION: AND METHODS OF USE THEREOF
: FILE REFERENCE: TRIPEP.003A
: CURRENT APPLICATION NUMBER: US/09/370,368
: CURRENT FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 510
: TYPE: PRT
: ORGANISM: Simian Immunodeficiency Virus
US-09-370-368-5

```

```

alignment_scores:
  Quality: 95.50      Length: 322
  Ratio: 0.632      Gaps: 24
  Percent Similarity: 46.894      Percent Identity: 23.913

```

```

alignment_block:
US-09-303-518d-125 x US-09-370-368-5 ..

```

```

Align seg 1/1 to: US-09-370-368-5 from: 1 to: 510

```

```

516 GGATTTCAAACCGCGCCTGTGTGATTGAGCCGCTTGGACCGACGCAAA 565
|||||..... 516
173 GlyPheGlnAlaLeuSerGlnGlyCysThrProTyrAspLeuAsnGln.M 189
|||||..... 173
566 TCCATGTTGT.....AAGCAGCTGCG.....GCAGAC 594
|||..... 566
189 eLeuAsnCysValGlyGlnHisGlnAlaAlaMetGlnIleLeuArgGlu 205
595 GTCCGCTGAAATGCTGCCACATCGAACA...CATGATTCGGCGG 641
|||..... 595
206 IleLeuAsnGlnGlnAlaAlaAspTrpAspValGlnHisProGlnProG 222
642 CCCGATCTCTGCCGCTTG.....AGTGCACGACATTCATT 679
|||||..... 642
222 yProLeuProAlaGlyGlnLeuArgGlnProSerGlySerAspIle.... 237
680 TCATCGAGCCGCTGCGCGAATAAACCGTG.....TGG 714
|||..... 680
238 .....AlaGlyThrThrSerThrValGlnGlnGlnIleGlnTrp 250
715 ACCATCAATTACAGATGATATTCATGCGCGCTTGTTCACAGAG 764
|||..... 715
251 MetHisArgGlnGlnAsnProIleProValGlyAsnIleTyr..... 264
765 CCGTTCGACACCGAGCGCGTATTGCCCTAGGTGTTCTCAAGTCACA 814
|||..... 765
265 .....ArgArgTrpIleGlnLeuGlyLeuGlnLysCysVal 276
815 AACCGCGCCTTTCGCTA..... 832
|||..... 815
277 ArgMetLysAsnProValAsnIleLeuAspIleLysGlnGlyProLysG 293
833 CGGTT.....TGGGTGCAAGATGCAATTAATCTGCGCGCG... 871
|||||..... 833
293 uProPheGlnSerTyrValAspArgPheTyrLysSerLeuArgAlaGln 310
872 .....AATGTTGACACAGCAACACCGCGATTT 901
|||||..... 872
310 LnaIaAspProAlaValLysAsnTrpMetThrGlnThr..... 322
902 CCGGTGCTGATTGACAGCGCGCGATTACACAGCGCGCGCATTTATTGTG 951
|||..... 902
323 ProLeu.....IleGlnAsnAlaAsnPro..... 330
952 GGAGCTACCAACAATCAGATTCCGTTATCGAAGAAGCGCGACGACAAGA 1001
|||..... 952
331 AspCysLysLeuValLeuLysGlyLeuGlyMetAsnProThrLeuGln 346
1002 GCTGTTCGGCTGGGTGCGCGAGCGGA..... 1031
|||..... 1002
347 GluMetLeuThrAlaCysGlnGlyValGlyGlyProGlyGlnLysAlaAr 363
1032 .....CAATTAATCCATCAGCGGTACCAACCTCGGCCATTTCCT. 1070
|||..... 1032
363 gLeuMetAlaGlnAlaLeuLysGlnAlaPheGlnProGlyProLeuProP 380
1071 .....GAAAAACAACCTCTCAAGTTCAACAC... 1097
|||..... 1071
380 heAlaAlaAlaGlnGlnGlnGlnValArgTyrValLysCysTrpAsnGly 396
1098 .....AGCCCTCAACGGCGCGCGACCGCGCATGTGTC 1129
|||..... 1098
397 GlyLysGlnGlyHisThrAlaLysGlnLysLysAlaProArgArgGlnG 413
1130 CGATTGTACTTACGAGCGCGTGATGCCCTTGATATCTGCGCACCGCTG 1179
|||..... 1130
413 yCysTrp.....LysCysGlyLysProGlyHis...G 423
1180 CTTTGGCGGATTTAATCGTCGCGATACCGACGCGCGAGCGATGGG 1229
|||..... 1180
423 lMetAlaLysCysProGlyArgGln.....ValGly 433

```


1230 TTGCTGGAATGGA.....CGAAGAAGACCTGC...TTTGGCAGCT 1270
 :::::::::::::::::::::
 434 PheLeuGlyPheGlyProTyrGlyLysLysProArgAspMet 450
 1271 TCGCTGGCC 1280
 :::::::::::::::::::::
 450 aGlnIlePro 453

seq_name: /cgn2_6/prodata/1/laa/5B_COMB.pep:US-08-467-822-21

seq_documentation block:

Sequence 21, Application US/08467822
 Patent No. 5843460
 GENERAL INFORMATION:
 APPLICANT: Labigne, Agnes
 APPLICANT: Sauerbaum, Sebastien
 APPLICANT: Ferrero, Richard L.
 APPLICANT: Thiberge, Jean-Michel
 TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 NUMBER OF SEQUENCES: 44
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Finegan, Henderson, Farbow, Garrett &
 ADDRESSEE: Dunner
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,822
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/447,177
 FILING DATE: 19-MAY-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/432,697
 FILING DATE: 02-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 569 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..569
 OTHER INFORMATION: /note= "URE B - FIGURE 3."

Alignment_scores:
 Quality: 95.00 Length: 425
 Ratio: 0.562 Gaps: 21
 Percent Similarity: 39.765 Percent Identity: 20.706

alignment_block:
 US-09-303-518D-125 x US-08-467-822-21

Align seg 1/1 to: US-08-467-822-21 from: 1 to: 569

205 ACTGCGCGCGCTTCAGCAAAATCGCGCGATTCACCGTGGCAAAAGCG 254
 :::::::::::::::::::::
 161 ThrGlyProAlaAspGlyThrAsnAlaThrThrIleThrProGlyArgAl 177
 255 CGTACTTCAGTCACTGCTGATTCGCGTTGAAGCAAGCAAGAAATCGACT 304
 :::::::::::::::::::::
 177 AsnLeuLysSerMetLeuArgAlaAla..... 186
 305 TTGAACGCTACGCACTGAA.....GCGCTGCAAACTTAAGC 342
 :::::::::::::::::::::
 187 ..GluGluTyrAlaMetAsnLeuGlyPheLeuAlaLysGlyAsnValSer 202
 343 GCGGAAGAACTGCGCGCAACCTGATCAATCGGTTG..... 381
 :::::::::::::::::::::
 203 TyrGluProSerLeuArgAspGlnIleGluAlaIleGlyPheLys 219
 382TGAGTGGCGTGGCAACCGCTTCAGCAAAATTC 418
 :::::::::::::::::::::
 219 sIleHisGluAspTyr..... 224
 419 CTGCGCTGATGCGGAGCGGTTGCCATCTTC.....GTCAATGCAATG 462
 :::::::::::::::::::::
 225GlySerThrProAlaAlaIleHisIscLysLeuAsnValAla 238
 463 GACACCAATCCGCTG...GCTGCCAGCCCTACGATTCATCAAGAAGC 509
 :::::::::::::::::::::
 239 AspGluTyrAspValGlnValAlaIleHisThrAspThrLeuAsnGluAl 255
 510 C.....GCCAGGATTTCAACGCGCGCTGTGATATGAGCGGTTGA 553
 :::::::::::::::::::::
 255 aGlyCysValGluAsp.....ThreGluAlaIleA 266
 554 CCGAAGCGCAAAATCCATGTTGTAAAGCAGCTGCGGACAGCTCCGCTCT 603
 :::::::::::::::::::::
 266 LeuGlyArgThrIleHisThrPheHisThrGluGlyAla..... 278
 604 GAAATGCTGCCACATGCAAAACATGAATTCGCGCGCGCGCTCCG 653
 :::::::::::::::::::::
 279GlyGlyLeuHisAlaP 284
 654 CGGTTGAGTGCGACGACATTCATTC...ATCGACCGCGTGGCGCGA 700
 :::::::::::::::::::::
 284 AspValIleLysMetAlaGlyIlePheAsnIleLeuProAlaSerThr 301
 701 ATAAACGCTG...TGACCATCAAT.....TATCAAGATGA 735
 :::::::::::::::::::::
 301 snProThrIleProPheThrLysAsnThrGluAlaGluHisMetAspMet 317
 736 ATTAACATGGCCGT.....TTGTT 755
 :::::::::::::::::::::
 318 LeuMetValCysHisHisLeuAspLysSerIleLysGluAspValGlnP 334
 756 TGCAACAGCGCGCTGTAACACGCGGATGTCGCTTGGGTCGGAAGATAGC 805
 :::::::::::::::::::::
 334 eAlaAspSerArgIleArgProGlnThrIleAlaAla.....GluAsp 349
 806 AAGTCAACAAACCGCGCTGTGCTACCGCTTGTGGGTCGGAAGATAGC 855
 :::::::::::::::::::::
 349 LysLeuHisAspMetGlyIlePheSerIleThrSerSerAspSerGlnAla 365
 856 CAAATTAAGTGGCGGCGAATGTT..... 879
 :::::::::::::::::::::
 366 MetGlyArgValGlyGluValIleThrArgThrThrGlnThrAlaAspLys 382
 880GACACAGACA 889
 :::::::::::::::::::::
 382 sAsnLysLysGluPheGlyArgLeuLysGluGluLysGlyAspAsnAspA 399

seq.documentation_block:
Sequence 21, Application US/08432697
Patent No. 6248330
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauterbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Flunegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000

```

alignment_scores:          Length: 425
    Quality: 95.00
    Ratio: 0.562          Gaps: 21
Percent Similarity: 39.765    Percent Identity: 20.706

alignment_block:
US-09-303-51BP-125 x US-08-432-697-21    ..
Align seg 1/1    to: US-08-432-697-21    from: 1    to: 569

```

```

205 ACTGCGCCGGCTTCAGGCACAAATGCCCGCATCCACCGGGAGCAACG 254
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
161 ThrGlyProIleAspGlyThrAsnAlaThrThrLeuPheProGlyArgAl 177
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
255 CGACTCTGTGATGACGCGATTGGCCGTGAAGCAACGACGAATCGAGT 304
    |||::|||::|||::|||::|||::|||::|||::|||::|||::
177 AsnLeuLysSerMetLeuAlaAla..... 186
    .....
305 TTGACGCTAACGACCTGAA.....CGCCTGGCAAATCTAAGC 342
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
187 ..GlucIuTygAlaMetAsnLeuGlyPheLeuAlaLysGlyAsnValSer 202
    .....
343 GCGGAGAAGTGCGCGCGAACCTGATCCAAATCGCGTTG..... 381
    |||::|||::|||::|||::|||::|||::|||::|||::|||::
203 TygIleProSerLeuArgAspGlnIleGluAlaGlyAlaIleGlyPheLy 219
    .....
382 .....TGACTGGCGCTGGCACCGCGCTGCTTCACAATTC 418
    .....
219 sIleHisGluAspTrp..... 224
    .....
419 CTGCGCTCGATGCCGACGCCGTTTGGCATCTTC....GTCATGGCATG 462
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
225 .....GlySerThrProAlaIleHisHisIscySLeuAsnValAla 238
    .....
463 GACACCAATCCGCT...GCTCCGACCCCTACGGTCATTATCAAGAGC 509
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
239 AspGluTyArgAlaGlnValAlaIleHisThrAspHrLeuAsnGluAl 255
    .....
510 C.....GCCGAGATTTAACAGCGCGCTGTGTGTTAGAGCGCTTGA 553
    |::|||::|||::|||::|||::|||::|||::|||::|||::|||::
255 agLyCyvalGluAsp.....ThrLeuGluAlaIleA 266
    .....
554 CCGGACGCCAAATACATCTTTGTGACGAGCTGGCGAGACGTGCCGCT 603
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
266 IagIyaGrthrIleHisThrPheHisThrGluGlyAla..... 278
    .....
604 GAATAATGCTCCAAATCATGAAACACATGAATTGGCGCGGCCCGCATCTGC 653
    .....
279 .....GlyGlyGlnHisAlaLpr 284
    .....
654 CGGTTTAGTGCGACACCATTCATTC..ATCGACGCGGTGGCGCGCA 700
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::
284 oAspValIleLysMetalagIylcylurheAsnIleLeuProLasertHra 301
    .....
701 ATTAACCGTG..TGACCATCATAT.....TATCAAGATGTA 735
    |||::|||::|||::|||::|||::|||::|||::|||::|||::

```

```

301 snProthrlleProPheThrLysasnThrGluAlaGluHisMetAspMet 317
736 ATTACATTTGGCCGT.....TTGTT 755
318 LeuMetValLysHisHisLeuAspLysSerIleLysGluAspValGlnPh 334
756 TGCACAGCGCGCTGTGACACAGCGCGGATTTGGCCGATTCGTC 805
334 eAlaAspSerArgIleArgProGlnThrIleAla.....GluAspG 349
806 AAGTCACAAACCGCGCTTCTTCGTAACGTTTGGTGCGAAAGTATCG 855
349 InLeuHisAspMetGlyIlePheSerIleThrSerSerAspSerGlnAla 365
856 CAATTAATCTGCGGCGCAATGGTT..... 879
366 MetGlyArgValGlyGluValIleThrArgThrTrpGlnThrAlaAspLy 382
880 .....GACACAGACA 889
382 sAsnLysLysGluPheGlyArgIleuLysGluGluLysGlyAspAsnAspA 399
890 AC.....CCGTGATTTCCGTCGTGATTTGAACGCGCGCAT 927
399 snPheArgIleLysArgTyrIleSerLysTyrThrIleAsnProGlyIle 415
928 ACACAGCGCGCGCATTTATTTGGACGCTACCAATCAGATTCGCT 977
416 AlaHisGlyIleSerAspTyrValGlySer..... 425
978 TATGAGAAGAGCGCGCAGCAAGAGCTGTTCGGCTG..... 1014
426 ValGluValGlyLysTyrAlaAspLeuValLeuThrPserProAlaPhe 442
1014 ..... 1014
442 heGlyIleLysProAsnMetIleIleLysGlyLysPheIleAlaLeuSer 458
1015 .....GTGCGCGCGCAGCGCGCAATA 1037
459 GlnMetGlyAspAlaAsnAlaSerIleProThrProGlnProValTyrTly 475
1038 CTCATCAGCGCTGACACCTCGCGCATTTCTGAAACCAACTCTTCA 1087
475 r.....ArgGlnMetPheGlyHisHisGlyLysasn.....L 486
1088 ACTTCACACAGCGCTCAAC..... 1107
486 ySPheAspThrAsnIleThrPheValSerGlnAlaAlaIatYrLysAlaGly 502
1108 .....GCGCGCGCAGCGCGCATTTGTCGCGATTTGTTACTTA 1142
503 IleLysGluGluLeuGlyLeuAspArgAlaAlaProProValLysAsnCy 519
1143 CGAGCGCGTGAATGCCCTTGATATC 1167
519 sArgAsnIleThrLysLysAspLeu 527

seq_name: /cgn2_6/prodata/1/1aa/6b_COMB.pep:us-08-466-248-21
seq_documentation_block:
; Sequence 21, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Faradow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: protein
; LOCATION: 1..569
; OTHER INFORMATION: /note="URE B - FIGURE 3."
; US-08-466-248-21

alignment_scores:
Quality: 95.00 Length: 425
Ratio: 0.562 Gaps: 21
Percent Similarity: 39.765 Percent Identity: 20.706

alignment_block:
US-09-303-518D-125 x US-08-466-248-21 ..
Align seg 1/1 to: US-08-466-248-21 from: 1 to: 569

205 ACTGCGCGCGCTTCAGCAAAATCGCGCGATTCACCGTGGCGAAAGCG 254
||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 ThrGlyProAlaAspGlyThrAsnAlaThrThrIleThrProGlyArgAl 177
255 CGTACTTCAGTCAGTGTGATTCGCCGTTGAAGCAACGCAAAATCGACT 304
||||:|||||:|||||:|||||:|||||:|||||:|||||:
177 aAsnLeuLysSerMetLeuArgAlaAla..... 186
305 TTGAACCTACGACACCTGAA.....GCGTGGCAAACTTAACG 342
||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ..GluGluTyrAlaMetAsnLeuGlyPheLeuValLysGlyAsnValSer 202
343 GCGCAAGAGTGGCGCGCATCGATCAATCGGTTG..... 381
||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 TyrGluProSerLeuArgAspGlnIleGluAlaGlyAlaIleGlyPheLy 219
382 .....TGACATTCGCGCTGCGCACCGCGTCGTTGACGAAAAATTC 418

```

```

219 silehISgluAspTrp..... 224
219 silehISgluAspTrp..... 224
419 STGCGCGTGCATGCCGAGCGCTTCCGCAATTC.....GTCAATCCGATG 462
225 .....GlySerThrProAlaAlaIleHisCysLeuAsnValAla 238
463 GACACCAATCCGCGT.....GTCGCCGACCTACGTCATATCAAGAGC 509
239 AspSerIuTyArgValGlnValAlaIleHisThrAspThrLeuAsnGluAl 255
510 C.....GCCGAGATTTCAAACGCGCGCTGTGATTCAGCCGTTGA 553
255 aglyCysValGluAsp.....ThrLeuGluAlaIleAla 266
554 CCGAAGCCAAATTCATGTTGTAGGACGTCGCGACAGCTGCCGCTC 603
266 IaglyArgThrIleHisThrPheHisThrGluGlyAla..... 278
604 GAAATGCTGCCAATCGAAGACATGATTCGCGCGCGCGCATCTCC 653
279 .....GlyGlyGlyHisAlaBr 284
654 CCGTTGAGTGCAGCAGCATTCATTC.....ATCGAGCGCGTGGCGCGCA 700
284 oAspValIleLysMetAlaGlyIuPheAsnIleLeuProAlaSerThrA 301
701 ATAAACCGTG...TGACCATCAT.....TATCAAGATGTA 735
301 snProThrIleProPheThrLysAsnThrGluAlaHisMetAspMet 317
736 ATTAACATTCGCGT.....TTGTT 755
318 LeuMetValCysHisHisLysAspLysSerIleLysGluAspValGlnPn 334
756 TGCAACAGCGCGTCTGAACACGCGCGTATGCTGATGCTGCTTC 805
334 eAlaAspSerArgIleArgProGlnThrIleAlaAla.....GluAspG 349
806 AAGTCACAAACCGCGCTCTGCGTACCGCTGTTGGTGGCGAAGATGCG 855
349 IuLeuHisAspMetGlyLePheSerIleThrSerSerAspSerGlnAla 365
856 CAATTAATCGCGCGCGATTCGTT..... 879
366 MetGlyArgValGlyIuValIleThrArgThrTrpGlnThrAlaAspLy 382
880 .....GACACAGACA 889
382 sAsnLysLysGluPheGlyArgLeuLysGluGluLysGlyAspAsnAspA 399
890 AC.....CGCGTATTCGCGTTCGGATTCGAACGCGCGCAT 927
399 snPheArgIleLysArgTyrlleSerLysTyrlleAsnProGlyIle 415
928 ACACAGAGCGCGCATATTCGAGCGCTACACCATCAGATTCGCT 977
416 AlaHisGlyIleSerAspTyrlleLysSer..... 425
978 TATCGAAGAGCGCGCGACGCAAGAGCTGTGCGCTG..... 1014
426 .ValGluValGlyLysTyrlleAspLeuValLeuThrSerProAlaPheP 442
1014 ..... 1014
442 heGlyIleLysProAsnMetIleLysGlyLysPheIleAlaLeuSer 458
1015 .....GTTCGCCGCGACCGCGCAATA 1037
459 GlnMetGlyAspAlaAsnAlaSerIleProThrProGlnProValTyTy 475
1038 STGCATCAGCGCTACACCGCTGCGCATTCCTGAAACAAACACTCTTCA 1087

```

```

475 r.....ArgGluMetPheGlyHisGlyLysAsn.....L 486
1088 AGTTCACACAGCGCGTCAAC..... 1107
486 yAspPheAspThrAsnIleThrPheValSerGlnAlaAlaTyrlleGly 502
1108 .....GGCGGCGACCGCGCATGTCGATTCGCTACTTA 1142
503 IleLysGluGluLeuGlyLeuAspArgAlaAlaProValLysAsnCy 519
1143 CGAGCGCGTATGCCCTGATATC 1167
519 sArgAsnIleThrLysLysAspLeu 527
seq_name: /cgn2_6/ptodata/1/aa/5B_COMB.pep:US-08-564-972-1
seq_documentation_block:
; Sequence 1, Application US/08564972
; Patent No. 5843462
; GENERAL INFORMATION:
; APPLICANT: Cont-Fine, B. M.
; TITLE OF INVENTION: DIPTHERIA TOXIM EPITOPES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,972
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.344051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-564-972-1

```

alignment_scores:

```

Quality: 94.50 Length: 489
Ratio: 0.450 Gaps: 28
Percent Similarity: 42.945 Percent Identity: 20.245

```

alignment_block:

```

US-09-303-518d-125 x US-08-564-972-1 ..
Align seg 1/1 to: US-08-564-972-1 from: 1 to: 535

```

```

10 ATCAAAAAGTCTAAACCTGCCATCCGGCGACAGCCGCT 59
   |||:|||||:|||||:|||||:|||||:|||||:
31 Ileglnysglyllelnlnysprolyserely..... Thrlnclnlys 45
60 TTACGAC.....:|||||:.....:66
45 ntyraspasprplysglyphetyrserThraspnllystraspa 62
67 .....GGCCGGCCATT...ACCAAGTCGGCTTGGCGAA..... 102
   |||:|||||:|||||:|||||:|||||:
62 laaiaaglytyrserValaspasnlnlnsnprouserglylvalagly 78
103 .....GAATAGCCGGTATCGGCC.....TCGATGAA 131
   |||:|||||:|||||:|||||:|||||:
79 GlyValVallyValThrtyrproglyleuThrlyValleuValleu 95
132 AGTCAGGAAGGCGATGCCCTCAAAAAAGCCAAAGTCGTTGAAGACA 181
   |||:|||||:|||||:|||||:|||||:
95 sValaspasnlaaglthrille.....:L 103
182 AAAAGAAATCCGGCGGTGTCTTACTGCGCGGCTTCAGGCAAAATCGCC 231
   |||:|||||:|||||:|||||:|||||:
103 ysllysglyleuThryrserleuThrGluProuleuetoGluValGly 119
232 GCG.....ATCACCGT.....GGCAAAAGCGGTACTTCA 263
   |||:|||||:|||||:|||||:|||||:
120 ThrGluGluPheIlelysarGpneGlyAspGlyAlaserArgValValle 136
264 GTCAGTCGATGATCCGTTGAGGCAAGC.....:294
   |||:|||||:|||||:|||||:|||||:
136 uSerleuProphela...GluIlyserSerSerValGluIryIleAsna 152
295 .....GAATCGAGTTTGA 309
   |||:|||||:|||||:|||||:|||||:
152 snTprGluGlnAlaLysAlaLeuSerValGluLeuGlnIleAsnheGln 168
310 .....CGCTACGCACTGAAGC 326
169 ThrArgGlyLysArgGlylnAspAlaMetGlyIuTyrmelaGlnAl 185
327 GCTGGCAAACTTAAGCGGGAAGAGTGGCGCAACTGATCCAAACCG 376
   |||:|||||:|||||:|||||:|||||:
185 acysala.....GlyAsnArgValArgArser...ValGlySers 198
377 GTTG.....TGACTGCGCTGCGCACCCGTCGCTC 408
   |||:|||||:|||||:|||||:|||||:
198 erleuSerGlyIleAsnLeuAspIrrpaspValIleArgAspIlyThrLys 214
409 ACCAAAATTCCTGCGGTGATGCCGAGCGCTTCGCATCTTCGTCATGC 458
   |||:|||||:|||||:|||||:|||||:
215 ThrLysIleGluSerleu.....LysGluHisGlyProIleLysasnLy 229
459 GATGACACCAATCCGCTGCGTCCGCGACCCATTCATTATCAAGAAG 508
   |||:|||||:|||||:|||||:|||||:
229 shetserGluSerProasnLysThrValSerGluGluLysAlaLysGln 246
509 CGCGCGAGATTTCAACGCGGCTGTGATTCAGCGCTTGAACGGA 558
   |||:|||||:|||||:|||||:|||||:
246 yrlleuGluGluPheHisGlnThrAlaLeuGlnHisProGluLeuSerGln 262
559 CGCAAAATCCATGTTTGAAGCAGCTGGCGCACAGCTGCCGTCGAAAA 608
   |||:|||||:|||||:|||||:|||||:
263 LeuLysThr.....:265
609 TGCTGCCAACATCGAAACACATGAATTGCGGCGCCGATCCTCGCGGT 658
266 .....:V 266
659 TGAATGGACGACATTCATTATTCAGAGCGGTCGCGGAAATTAAC 708
   |||:|||||:|||||:|||||:|||||:
266 alThrGlyThrAsnProValPhe.....AlaGlyAlaAsnTyrla 279
709 GGTGACCAATCAATTATCAAGATGTAATACC..... 741
   |||:|||||:|||||:|||||:|||||:
280 AlaIrrpAlaValAsnValAlaGlnValIleAspserGluThrAlaAsps 296
742 .....ATTGGCGGT 751
   |||:|||||:|||||:|||||:|||||:
296 nleuGluLysThrThrAlaAlaLeuSerIleLeuProGlyIleGlySer 313
752 TG.....TTTGACACAGCGCTCTG.....AACACGAGCGCGTAT 789
   |||:|||||:|||||:|||||:|||||:
313 alMetGlyIleAlaAspGlyAlaValAlaHisAsnThrGluGlnIleVal 329
790 GCCCTAGGTGTTCTCAAGTCACAAACCGCGCTTCGTGACGCTT 839
   |||:|||||:|||||:|||||:|||||:
330 Ala.....GlnSerIleAlaLeuSerSerleu 339
840 GGGTGCACAAATATCCAAATTAATCTCGGCGCAATGTTGACACAGACA 889
   |||:|||||:|||||:|||||:|||||:
339 lValAlaGlnAlaIleProleu...ValGlyGluLeuValasp..... 352
890 ACCGCGTATTCGCGTTCGATTCAGAGGCGCATTCACAGCGCG 939
   |||:|||||:|||||:|||||:|||||:
353 .....IleGlyPheAlaAla 357
940 CAGCATTAATTTGGACGCTACCAATCAGATTCGTTATCAGAGAAG 989
   |||:|||||:|||||:|||||:|||||:
358 TyraspPheValGluSerIleIleAsnLeuPheGlnValAlaHisAsn 374
990 CCCGACCAAGAGCTGTCGGCTGCGGCTGCGG.....CAGC 1027
   |||:|||||:|||||:|||||:|||||:
374 rTyraAsnArgProAlaTyf.....SerProLysIlystrGlnP 388
1028 CG.....GACAAATCTCATCAGCGGTACACCTGCGCATTTTC 1068
   |||:|||||:|||||:|||||:|||||:
388 roPheLeuHisAspGlyTyrlaValaserTrrpAsnThrValGluAspser 404
1069 CTGAAAAACAACCTCTCAAGTTCAACACAGCCGTCACAGCGCGCACCG 1118
   |||:|||||:|||||:|||||:|||||:
405 Ile.....IleArgThrGlyPheGlnGlyIuSerG 415
1119 CGCGATGTCGCGATTTGACTTACAGAGCGGTGATCGCTTGAT.... 1164
   |||:|||||:|||||:|||||:|||||:
415 yHisAspIleLysIleThrAlaGluAsnThrProleuProIleAlaLy 432
1165 ..ATCTGCCACCTG 1179
432 alLeuLeuProThrIle 437
seq_name: /cgn2_6/plodata/1/laa/6B_COMB.pep:us-09-171-969-9
seq_documentation_block:
: Sequence 9, Application US/09171969
: Patent No. 6284533
: GENERAL INFORMATION:
: APPLICANT: Thomas, Lawrence J.
: TITLE OF INVENTION: PLASMID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Banner & Witcoff, Ltd.
: STREET: 75 State Street, Suite 2300
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1807
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Wordperfect 6.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/171,969
: FILING DATE: 01 May 1997 (01.05.97)
: CLASSIFICATION: 514

```

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/640,713
 FILING DATE: 01 May 1996 (01.05.96)
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/802,967
 FILING DATE: 21 February 1997 (21.02.97)
 ATTORNEY/AGENT INFORMATION:
 NAME: Leon R. Yankwich
 REGISTRATION NUMBER: 30,237
 REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 535 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL:
 ANTI-SENSE:
 FEATURE: (A)NAME:
 LOCATION:
 US-09-171-969-9

alignment_scores:
 Quality: 94.50 Length: 489
 Ratio: 0.450 Gaps: 28
 Percent Similarity: 42.945 Percent Identity: 20.245

alignment_block:
 US-09-303-518D-125 x US-09-171-969-9 ..

Align seg 1/1 to: US-09-171-969-9 from: 1 to: 535

```

10 ATCAAAAGCTTAACCTGCCATCGCGGAGACCGAGACCGCT 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
31 Ileglnysgylleglnysprolysseryl.....Thrglnlyas 45
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
60 TTACGAC..... 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
45 ntyraspsasprtrprrysglyrhetryserthrinspsnlystyrasp 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
67 .....GGCCGCGCATTT...ACGAGCTCGCTGCTGGCGAA..... 102
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
62 laalaglytyrservalaspasnlglnsnproulserglylsalagly 78
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
103 .....GAATATGCCGTATGCCGCC.....TCGATGAA 131
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
79 glyvalvalylvalthrtyrproglyneuthrlystalleualaleuyl 95
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
132 AGTCAAGAGAGCGATGCCGTCAAAAAGCCAAAGCTCTTTGAGACA 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
95 svalaspsasnlaglthrile.....L 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
187 AAAAATCCGGCGCTGTGTACTGCGCGGCTTCAGGGAATAGCC 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
103 yslgsluendglyleuaserleuthrpgluolemetgluInvalagly 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
120 thrngluuphellelyargpheelyspglylaserargvalvalle 136
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
264 GTCAGTCGTGATTCGCTTGAAGCAAGAC..... 294
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
136 uSerleupropheala...GluGlyserSerSerValglutrylleasna 152
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
295 .....GAATGAGCTTTGAA 309
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
152 sntrpgluInlalysalaleuSerValgluendglulleasnphgln 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
310 .....CGCTACGACCTGAGC 326
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
169 ThrargglyllysargglylnaspalameTyrglutryMetalaInal 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

327 GCTGGCAAACTTAAGCGGCGAAGATGCGCGGCAACCTGATTCATCCG 376
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
185 acysala.....GlyasnargvalargSer...Valglyers 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
377 GTTTG.....TGGACTGGCGCGGACCGCGCGCTTC 408
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
198 erleuSerCyslleasnleuasprrpasrvalilleayasppllyThrLys 214
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
409 AGCAAAATTCCTCGCGCTGATCCGCGCGGCGGATTCCTGCAATGC 458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
215 ThrlyslleuSerleu.....Lysglunhsglyproilleysanly 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
459 GATGACACCAATCCGCTGCGCTGCCGACCTACGCTCATTTCAAAAG 508
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
229 smetSerGluSerProasnlyThrValsergluInlulysalalyserInt 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
509 CCGCGGAGATTTCAACCGCGCTGTGTGATGAGCGTTTGACGCA 558
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
246 yrluendgluphehlsnglnthrAlaleuInhlsProgluendserclu 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
559 CGCAAAATCCATGTTTGTAAAGCAGCTGGCGAGACGTGCGCTGA 608
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
263 LeulysThr..... 265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
609 TGCTGCCAATCATGAACACATGATTCGGCGCGCGCATCCTGCGGTT 658
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
266 .....V 266
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
659 TGAATGACACGACATTCATTCATGAGCGCGCGCGAATTAAC 708
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
266 althrghlythrpsnprovalphe.....AlaglylalsnlyrAla 279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
709 GTGTGACATCAATTAATCAAGATTAATACC..... 741
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
280 AlatrplavalasnlvalalaglnvalilleaspsergluthrAlasps 296
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
742 .....ATTGCCGCTT 751
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
296 nleuInlysthrThrAlalaleuSerlleuProglylleglyserV 313
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
752 TG.....TTTGCACAGCGCGCTGTG.....AACCCGACGCGCTGAT 789
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
313 almetglyllealaspglyAlalavlnhlsasnthrIngluInleval 329
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
790 GCCCTAGCTGTCTCAAGTCACAAACCGCGCGCTTGCAGTACGCTTT 839
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
330 Ala.....GlnserllealeuSerleu 339
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
840 GGTGCGAAAGTATCGCAATTAATCTGCGGCGAATGTTGACACAGACA 889
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
339 tValAlaglnAlalThrleu...ValglyluInleuValasp..... 352
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
890 ACCCGGTGATTTCCGGTTCGTATGAACGCGCGGATTCACACAGCGCG 939
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
353 .....lleglyrhealiala 357
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
940 CACGATTTTGGAGCCTACACACATCAATTCGCTTATTCAGACAGAG 989
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
358 tyrasnphelvalgluserllelleasnleupheglInvalValhlsans 374
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
990 CCGCAGCAAGAGCTGTTCGGCTGGGTGGCGCG.....CAGC 1027
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
374 rTytnasnargProAlatyr.....SerProglunhlsLysThrInp 388
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1028 CG.....GACAAATCTCATCGAGGCTCAACACCTCGGCGCATTTTC 1068
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
388 toPheleuInhlsaspLytyrAlalaserlprasnThrValgluasps 404
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1069 CTGAAAAAACAACCTTCAGAGTTCACACAGCGCTCAGCGCGGACCG 1118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
405 lle.....lleargthrcllypheelnglylusergl 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
1119 CGCCATGTGCGGATGTGTACTTACGAGCGGCTATGCCCTTGGAT... 1164
   |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

415 YHISAPLEIETHEHIALAGLUNSNTHPROLEUPTHELEAGLYV 432
1165 ..ATCCGCCACCCCTG 1179
432 alteleuPTHLE 437

seq_name: /cgn2_6/ptodata/1/laa/6B.COMB.pep:us-09-081-345-18

seq_documentation block:

Sequence 18, Application US/09081345

Patent No. 6228641

GENERAL INFORMATION:

APPLICANT: Bahija Jallal

TITLE OF INVENTION: Gregory D. Plovman

TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/081,345

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/047,222

FILING DATE: May 20, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 234/253

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 802 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-081-345-18

alignment_scores:

Quality: 94.50

Ratio: 0.559

Percent Similarity: 51.057

Percent Identity: 22.356

alignment_block:

US-09-303-518D-125/rev x US-09-081-345-18

Align seq 1/1 to: US-09-081-345-18 from: 1 to: 802

942 GTGGGCGCTGTGTATCGCGCGCTTCAATACCGAAGAAATCAGC 893
306 IIEIIMIAIAGIINYSERIEPIROGLINISERLEUTHR...VALGIUNL 321
892 GGTGTCTGTGTCAACCAATTCGCCGCGAGTATTTGGCATCTTTCGA 843

321 ASPSERCYSPROLEUASPLEUPROLYSASNALAMETALGASPVALLYST 338
842 CCCAAACGGGTACGCAAGAGCGGGTGTGTGACTTGTGAAACCACTAG 793
338 HTHTASH...GLNHISERLYSGIINGIYALAGIUNLAIUSERTHR... 352
792 GGCATATCAGCGGCTCGGTGTTCAGACGGCTGTGCAAAACAGGCCA 743
353 GLYLYSERSEULEUGLYLEUARGYTHSER...THMETAS 365
742 TGTATTTACATCTTGATATTTGATGTCACACGGTTTATTCGGCGC 693
365 NALAGIUNGLIUNLEUVALLEHISERIALYSSERSEPROSERPHEA 382
692 ACCGGCTCGATGAATGTCGCGCC...ACT 661
382 SNCYSLIU...GLIDEASNICYSGLYCYASNASNLYSALAVALLIETHR 397
660 CAACCGGCGAGATCGCGCGCC...GA 635
398 ARGASNGLYGINALARGALASERPROVALVALGYLUPROLEUINLY 414
634 ATTCATGTGTTCGATGTGCGACATTTTCAGACGACGTCGCGCA 585
414 STYGLINISERLEUASPPHEGLYSERMETLEU...PHYLY...SERCYSPRO 429
584 GCTGCTTACAAACATGATTTGCGTTCGCAACGCGCATACCA 535
430 SERALALEUPRO...ILEASNHRILASAPHYLYTHIASNSE 443
534 CAGG...CCGCGTTGAAATCTCGCGGCTTGTGATA... 498
443 RYSGIYPROVALYLSARGYTHLYSERTHRPHEGLIUNLEUING 460
497ATGACCGGAGGTGCGACGACGCGGATGTG 465
460 LARGLYSTRASNSPLEUNLALVALGYLSPGLYPHESECYSLIUN 476
464 TCC.....ATCCGATTCAGCAAGAT 445
477 SERGLIUNHISGLIUNISTYRSEULEUARGGLIUNGLIUNLALINARGVA 493
444 GCGCAACGGCTCGGATGATGACGAGGAGGATTTGCTCAACGAGCGGTG 395
493 IALAHISVALSERSEGLIUNLEUASNTRYRSEULEUPROGLYALCYSA 510
394 GCAGCGCATC.....CACAAACCGGATTCAGTACGTCGCGCG 354
510 SPALASERCYSVALPROARGHISERPROGLYALALAUARGVALHISLEU 526
353 ACTTCCTGCGCGCTTAAGTTTGCAGCGCTTCAGTGCCTGCA 304
527 TYRTHRISERLEUALAGIUNASPROTYRPHESER..... 537
303 CTCGATTTGCTGTGCTTCAACGCGCATCAGACTGACGATGACG 254
538SERSEPROASNSERALA.....ASPSELYSMETS 549
253 GCTTTTCGCGACGCGTGAATGCGCGCATTTTGCCTTAACCGCGCGCAGTA 204
549 ERPHASP.....LEUPROGLIUNYSGIUNASPLY 558
203 AACACACGCGCGGATTTTGTCTCAACAGACTTGCGCTTTT 154
559 ALATHISERPROGLYALALAULEUPROALASERTHTRHSEPHNEP 575
153 GACGCGATCGCTTCTTCAATGAGGGCGCATACCGCATAT 104
575 ETRYSERASNPROIHASPSERLEUVALMETASNHRILEUTHRSEPHES 592
103 CTTCGCCAAGCAACGCGACTCGTAAATGCGCGCGCGCTCG 63
592 ERPROLEUASNGIUNGLIUNHIALVALGIALAIPROSER 605

seq_name: /cgn2_6/plodata/1/aa/5B_COMB.pep:US-08-760-489-2

seq_documentation_block:

Sequence 2, Application US/08760489

Patent No. 5830696

GENERAL INFORMATION:

APPLICANT: Short, Jay M.

TITLE OF INVENTION: DIRECTED EVOLUTION OF THERMOPHILIC

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson, P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/760,489

FILING DATE: 05-DEC-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/008,311

FILING DATE: 07-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Ph.D., Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 09010/008001

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1487 amino acids

TYPE: amino acid

TOPOLOGY: 1linear

MOLECULE TYPE: protein

FRAGMENT TYPE: Internal

US-08-760-489-2

alignment_scores:

Quality: 94.50 Length: 547

Ratio: 0.475 Gaps: 27

Percent Similarity: 36.380 Percent Identity: 20.658

Alignment_block:

US-09-303-518D-125 x US-08-760-489-2 ..

Align seg 1/1 to: US-08-760-489-2 from: 1 to: 1487

```

15 AAAGGTAAACCTGCGGCGGAGCGGAGCGGAGCGGCGTTACG 64
   :::::::::::::::::::: |||:::
295 ArgGlnArgProAlaArgGlyGlyAspGlnLeuArgThrLeuAla 311
   :::::::::::::::::::: |||:::
65 ACGG.....CCCGGCATTACCGGAAGTCGGTTGCT.....T 96
   |||::: |||::: |||:::
311 GGIuGlySerProArgGlnArgProArgSerGlnArgLeuGlnAsp 328
   |||::: |||::: |||:::
97 GCGGAAGATATGCGCGTATGCGCCCGCGATGAAGTCAGGAAGCGA 146
   |||::: |||::: |||:::
328 rPasnArgVal..... 331
147 TCGCGTCAAAAAAGCGCAAGTCGTGTGAAGACAAAAAGATCCGGGCG 196
   :::: |||::: |||:::
332 .....GluGlnAspLeuSerLeuAl 338

```

```

197 TGCTGTTACTGCGCGCGC.....TTCA 219
   :::: |||::: |||:::
338 AsnValValLysGlyGlyArgAlaGlyGlnLeuAlaGlyThrArgGlyVal 355
   :::::::::::::::::::: |||:::
220 GCGAATATCGCGCGGATTCACCGTGGCAAAAGCGCGTACTTCAGTCAGT 269
   |||::: |||::: |||:::
355 rGlnAsnArgArgHisAla.....ArgArgAlaArg..... 365
   |||::: |||::: |||:::
270 CGTATTCGCGTGAAGCGAAGCGAATCGATTGTAACGCTACGCGAC 319
   |||::: |||::: |||:::
366 ArgAspSerGlu..... 369
   |||::: |||::: |||:::
320 CTGAAGCGCGTGGCAAA.....CTTAAGCGCGAAGATCGCGCGCA. 362
   |||::: |||::: |||:::
370 ....SerSerGlyAspSerLeuLeuProArgTyrArgAlaProGln 385
   |||::: |||::: |||:::
363 .....CCTGATCCCAATCCGCTTTGTGCGAC 386
   |||::: |||::: |||:::
385 LValGlyLeuGlnGlnHisArgGluProGlnProLeuHisAlaProPro 401
   |||::: |||::: |||:::
387 TCGCGTGGCGCGACCG..... 401
   |||::: |||::: |||:::
402 LeuAlaSerArgSerAspAsnArgGluGlyGluGlyProHisGlnArgAs 418
   |||::: |||::: |||:::
402 .....TCGCG 405
   |||::: |||::: |||:::
418 pThrLeuGlyArgAlaGlyGlnArgGlyGlyValArgGlnValArgGly 435
   |||::: |||::: |||:::
406 TTCAGCAAAATTCCTGCGCGTGCATCGCGCGGCTGCGCATTCGTCGCA 455
   |||::: |||::: |||:::
435 aHisArgGlnArgThrArgGlyProArgTyrValGlnHisLeuGlnArg 451
   |||::: |||::: |||:::
456 TCGGATGAGCAC.....CA 469
   |||::: |||::: |||:::
452 AlaAspGlyArgGlyGlyAlaArgLeuProArgAlaLeuLeuArgLeuSe 468
   |||::: |||::: |||:::
470 ATCGCGTGGCGCGCGCGCGTACGGT.....CAT 498
   |||::: |||::: |||:::
468 rAlaGlyGlyTyrGluProArgGlyGlyGlyAlaGlnAspProGlnHisA 485
   |||::: |||::: |||:::
499 ATCAAGAGAGCGCG..... 512
   |||::: |||::: |||:::
485 sPlysArgProArgThrGlnGlnAspAspGlyGlyValArgGlnGly 501
   |||::: |||::: |||:::
513 .....CGAGATTTCGAAGCGCGCGTGGATGAGCGCGTTGAGCG 556
   |||::: |||::: |||:::
502 LysGlyArgGlyPheProLeuArgGlyArgGlyArgAspAsnLeuGln 518
   |||::: |||::: |||:::
557 AACGCAAAATTCATGTTTGAAGCAGCTGGCGCAGCGTGGCGTCTGAA 606
   |||::: |||::: |||:::
518 nHisArgArg...CysLeuSerLeuArgGlnArgProLysGlyArgG 534
   |||::: |||::: |||:::
607 AATGTCGCAACATCGAAGACACATGAATTCGGCGCGCGCGATCTGCGG 656
   |||::: |||::: |||:::
534 LysGlyArgGlyArgGlnLeu...LeuProGlnArgAlaLeuLeuArg 549
   |||::: |||::: |||:::
657 TTTGAGTGGCGCAGCACATTCATTTCATCGAGCC..... 689
   |||::: |||::: |||:::
550 ArgAsnProGlnGlnAlaGlnHisArgValArgArgAspLeuArgG 566
   |||::: |||::: |||:::
690 .....GCTGCGCGCGAATGAACCGGTGCGACCATCATCAT 723
   |||::: |||::: |||:::
566 nSerSerAlaSerGlnGlyGlnArgLeuAspArgGlyLeuHisGlnAla 583
   |||::: |||::: |||:::
724 TATCAAGATGTAATTAC.....CATTCGCGCGTTGTT 755
   |||::: |||::: |||:::
583 rGSerArgGlnValPheGlyAlaGlnValProGlnHisThrProAspLe 599
   |||::: |||::: |||:::
756 TGCACACAGCGCGTCTGAACACGAGCGCGTGATTCGCGTGGTGTCTC 805
   |||::: |||::: |||:::
600 LeuProGlySerSerGlnLeuArgLeuArg..... 609
   |||::: |||::: |||:::
806 AAGTCAACAACCGCGCGCTTGGGTACCGTTTGGGTGCGAAGATATCG 855

```

[illegible]

```

1206 TACGACAGCCCGCAGCATTTGGCTTGGTAATTGACGACGAGA..... 1250
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
679 Glyceratgagnginglyleuproleinglylaalasparglnleuhr 695
1251 ....AGACCTGGCTTTGGAGCTTGCTGTGCCCGGGCAATATGCAATAC 1296
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
695 gvaiglyprrarphe.....Glnasplguvala 705
1297 GAGCCGCGTGTGGCAAAAGTCTGGAAACCATTTGAAAGAGA 1337
      |||||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
705 rgproleu.....SerGlySerHisasnIngly 714

seq_name: /cogn2_6/prodata/1/1aa/6B_COMB.pap:US-09-185-373-2

seq_documentation_block:
  sequence No, Application US/09185373
  Patent No. 6335179
  GENERAL INFORMATION:
    APPLICANT: Short, Jay M.
    TITLE OF INVENTION: DIRECTED EVOLUTION OF THERMOPHILIC
      ENZYMES
    NUMBER OF SEQUENCES: 4
    CORRESPONDENCE ADDRESS:
      ADDRESSSEE: Fish & Richardson, P.C.
      STREET: 4225 Executive Square, Suite 1400
      CITY: La Jolla
      STATE: CA
      COUNTRY: US
      ZIP: 92037
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette
      COMPUTER: IBM Compatible
      OPERATING SYSTEM: Windows95
      SOFTWARE: FastSeq for Windows Version 2.0
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/185,373
      FILING DATE: 03-NO. 6335179-1998
      CLASSIFICATION: <unknown>
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US/08/760,489
      FILING DATE: 05-DEC-1996
      APPLICATION NUMBER: 60/008,311
      FILING DATE: 07-DEC-1995
    ATTORNEY/AGENT INFORMATION:
      NAME: Haile, Ph.D., Lisa A
      REGISTRATION NUMBER: 38,347
      REFERENCE/DOCKET NUMBER: 09010/008001
      TELECOMMUNICATION INFORMATION:
        TELEPHONE: 619/678-5070
        TELEFAX: 619/678-5099
    INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 1487 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
        FRAGMENT TYPE: internal
      SEQUENCE DESCRIPTION: SEQ ID NO: 2:
      US-09-185-373-2

alignment_scores:
  Quality: 94.50      Length: 547
  Ratio: 0.475      Gaps: 27
  Percent Similarity: 36.380      Percent Identity: 20.658

alignment_block:
  US-09-303-518D-125 x US-09-185-373-2 ..
  Align seg 1/1 to: US-09-185-373-2 from: 1 to: 1487

15 AAAAGCTTAAACTGCCCATCGGGGCAACCGAGACGCGCTTAGC 64
      ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||

```

```
295 ArgGlnArgArgProAlaArgGlyLysPlySglnLeuArgThrLeuArg 311
65 ACG.....CCCGCATTAACGAGTGGCT.....T 96
   ||| ||| ||| |||
311 glnGlySerProArgGlnArgProArgSerGlnArgLeuGlnAsp 328
97 GCCGAGAAATATGCGGTATGCGCCCTCGATGAAGTCAAGAGAGCGA 146
   ||| ||| ||| |||
328 rPasnArgVal..... 331
147 TCCCGTCAAAAAGGCCAAGTCTGTTTGAAGACAAAAGATCCGGCG.196
   ||| ||| ||| |||
332 .....GlnGlnAspLeuSerLeuAl 338
197 TGGTGTATTACTGCGCGCGC.....TTCA 219
   ||| ||| ||| |||
338 aasnValValCysGlyArgAlaGlnLeuArgThrArgGlnGlyA 355
220 GGCAAAATCGCGCGATTCACCGTGGGAAAGCGGCTACTTCAGTCACT 269
   ||| ||| ||| |||
355 rGlnAsnArgArgHisAla.....ArgArgAlaArg..... 365
270 CGTATTGCGGTTGAAGCAAGCAAGCAATCGAGTTTGAACGCTACGCAC 319
   ||| ||| ||| |||
366 ArgAspSerGlu..... 369
320 CTGAAGCGCTGGCACA.....CTTAAGCGCGAAGAGTGGCGCGCA. 362
   ||| ||| ||| |||
370 .....SerSerGlyAspSerLeuSerProProArgGlyArgAlaProGlnG 385
363 .....CTGATCCAATCGGTTTGTGAC 386
385 lylalGlyLeuGlnGlyHisArgGlyProGlnProLeuHisAlaProPro 401
   ||| ||| ||| |||
387 TGGCTGGCGCACCG..... 401
402 LeuAlaSerArgSerAspAsnArgGlnGlyGlnGlyProHisGlnArgAs 418
402 .....TCCG 405
418 pThrLeuGlyArgAlaGlyLysArgGlyLysAlaArgGlnValArgGly 435
406 TTGAGCAAAATTCCTGCGTCATGCGGCGGCTGGCCATCTTGCTGACA 455
   ||| ||| ||| |||
435 aHisArgGlnArgThrArgGlyProArgGlyValGlnHisLeuGlnArg 451
456 TGGGATGCGAC.....CA 469
452 AlaAspGlyArgCysGlyAlaArgLeuProArgAlaLeuArgLeuSe 468
470 ATCCGCTGCTGCGCGACCTACGGT.....CATT 498
   ||| ||| ||| |||
468 rAlaGlyGlyArgGlnProArgGlyGlyLysAlaGlyAsnProGlnHisA 485
499 ATCAAGAAGACCGC..... 512
485 sPlyAspArgProArgThrGlyLeuGlnAspAspGlyValAlaArgGlnGly 501
513 .....CGAGGATTCAAACGCGCGCTGTTGTTAAGCCGTTTGAACG 556
   ||| ||| ||| |||
502 LysGlyArgGlyPheProLeuArgGlyArgGlyArgAspAsnLeuGlnG 518
557 AACGCAAAATCCATGTTTGAAGCAGCTGCGGAGAGCGCGCTGAA 606
   ||| ||| ||| |||
518 nHisArgArg...CysLeuSerIleArgLeuGlnArgProLysGlyArgG 534
607 AATGCTGCCAATCGAAACACATGAATTCGCGCGCGCTGCTGCGCG 656
   ||| ||| ||| |||
534 luserCysArgLysArgGlnLeu...LeuProGlnArgAlaLeuLeuArg 549
657 TTGAGTGGCAGCAGCATTCATTTCATCGAGCC..... 669
   ||| ||| ||| |||
550 ArgAsnProGlnGlyGlnAlaGlnHisArgValArgArgAspLeuArgG 566
690 .....GCTCGCGCGAATAAACCGTGTGACCATCAT 723
   ||| ||| ||| |||
566 nSerSerAlaSerGlnGlyAlaArgLeuAspArgValArgLeuHisGlnA 583
724 TATCAAGATGAATTAC.....CATTGCGCGTTGTT 755
   ||| ||| ||| |||
583 rSerArgGlnValPheGlyAlaGlnValProGlnHisThrProAspIle 599
756 TGCAAAGCGCGCTGAAACACCGAGCGGTGATTCGCTAGTGTTCTC 805
   ||| ||| ||| |||
600 LeuProGlySerSerGlnLeuArgLeuArg..... 609
806 AAGTCAACAAACCGCGCTTGGGTACCGTTTGGGTGCGAAAGTATGCG 855
609 ..... 609
856 CAATTCATCGCGCGCAATTGGTTGACACAGACACACCGGTGATTTCCGG 905
610 .....LeuGlnAlaArgLysPhe... 615
906 TTGCTATTGACGGCGCATTTACACAGCGCGCGACATTTATTGGCAC 955
   ||| ||| ||| |||
616 .....PheArgArgArgLysAlaArgLysArgHisArgLeu...GlyAs 629
956 GCTACCAATTCAGATTTCCGTTATCGAAGAGCGCGCAAGAGCTG 1005
   ||| ||| ||| |||
629 rLeuSerGlyGlyAspLeuArgLeuAspLysArgGlyGlnGln.....I 644
1006 TTGCGGCTGGGTTGCGCGCGACCGGCAAAATCTCCATCAGCGGTAAAC 1055
644 LeuArg.....Gly 646
1056 CCTCGGCATTTCTGAAAACAACTCTTCAAGTTCAACACAGCGCTCA 1105
647 Pro.....GlyLeuArgHisArgLysAr 654
1106 ACGCGCGCGACCGCGCATGTCGCGATGTACTTACAGCGCGTGATG 1155
654 gaAsnSerArgPheAsnHis.....ProA 662
1156 CCTTGATATCTGCGCCACCTGCTTTTGGCGGATTTAATCGTCGGCA 1205
   ||| ||| ||| |||
662 lAlaValLeuProArgGlnProCysSerGlyAspGlyValArgGly 678
1206 TACCGACAGCGCGACGATGCTGCTTGAATTTGACGAGAGA..... 1250
   ||| ||| ||| |||
679 GlyLeuArgArgGlnGlyLeuProLeuGlnGlyAlaAspArgGlnLeuArg 695
1251 .....AGACCTGCTTTGTGACAGCTTCGTCGCGCGCAATACGAATAC 1296
   ||| ||| ||| |||
695 gValGlyProArgPhe.....GlnAspGlyValA 705
1297 GCGCGCTGTTGCGCAAGTGTGAAACCAATGAGAGCA 1337
705 rGProLeu.....SerGlySerHisAsnGlnGlnGly 714
seq_name: /cgnr_6/prodata/1/1aa/6B_COMB.pep:US-09-185-373-4
seq_documentation_block:
: Sequence 4, Application US/09185373
: Patent No. 6335179
: GENERAL INFORMATION:
: APPLICANT: Short, Jay M.
: TITLE OF INVENTION: DIRECTED EVOLUTION OF THERMOPHILIC
: ENZYMES
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
```

ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/185, 373
FILING DATE: 03-No. 6335179-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/760, 489
FILING DATE: 05-DEC-1996
APPLICATION NUMBER: 60/008, 311
FILING DATE: 07-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: HALL, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/008001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1487 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-185-373-4

alignment_scores:
Quality: 94.50 Length: 547
Ratio: 0.475 Gaps: 27
Percent Similarity: 36.380 Percent Identity: 20.658

alignment_block:

US-09-303-518d-125 x US-09-185-373-4 ..

Align seg 1/1 to: US-09-185-373-4 from: 1 to: 1487

15 AAAAGTCTAAACCTGCCATCGCGGCGACGCGAGCAAGCCGTTACG 64
:::|||||:::|||||:::|||||
295 ATGGlnAArgProAlaArgGlyLysPylsGlnLeuAArgThrLeuA 311
65 ACGG.....CCGGCCATTACCGAAGTCGGCTTGTCT.....T 96
:::|||||:::|||||:::|||||
311 gGlnGlySerProAlaArgGlnAArgProAArgSerGlnAArgLeuA 328
97 GCGGAAGATATGCGCGATATGCGCCCTCGATGAAGTCAAGAAAGCGGA 146
||:|||||:::|||||
328 rpsnAArgVal.....GlnGlnAspLeuSerLeuA 331
147 TGGCGTCAAAAAGCGCAAGTCGTGTTGAAGCAAAAAGAATCCGGCG 196
:::|||||:::|||||
332GlnGlnAspLeuSerLeuA 338
197 TGGTGTCTACTGCGCGGC.....TTCA 219
:::|||||:::|||||
338 aaenValValCysGlyArgAlaGlyGlnLeuAArgThrAArgGlnGly 355
220 GCGAAATGCGCGCGATTCACCGTGGCGAAAGCGGTAAGTCAAGTCA 269
|||||:::|||||
355 rglInAsnAArgGlnAArg.....ArgAArgAArg..... 365
270 CGTGAATGCGCGTGAAGCAAGCAAGCAATGAGTTGAAGCCTACGAC 319
|||||:::|||||
366 ArgAspSerGlu.....CTTAAGCGCGAAGCAAGTGGCGCGCA 362
|||||:::|||||

370SerSerGlyAspSerLeuProAArgTyrAArgAlaProGln 385
363CCTGATTCGAATCCGGTGTGGAC 386
385 lYAlaGlyLeuGlnGlnHisAArgGlnProGlnProLeuHisAlaProPro 401
387 TGGCGCGCGCACCG..... 401
402 LeuAlaSerArgSerAspAsnAArgGlnGlnGlyProHisGlnAArg 418
402TCCG 405
418 pTrpLeuGlyAArgAlaGlyLysAArgGlyValAArgGlnValAArgGly 435
406 TTGAGCAAAATTCCTGCGTCGATCGCGCGCGTGGCATCTTCGTCAA 455
||:|||||:::|||||
435 alHisAArgGlnAArgThrAArgGlyProAArgTyrValGlnHisLeuAArg 451
456 TCGATGAGCAC.....CA 469
452 AlaAspGlyAArgCysGlyAlaAArgLeuProAArgAlaLeuAArgLeu 468
470 ATCGCGTGGCGCGCGACCGTACGCT.....CAT 498
468 rAlaGlyGlyTyrGlnProAArgGlyLysAlaGlyAsnProGlnHisA 485
499 ATCAAGAGACCGC..... 512
485 sPylsAArgProAArgThrGlnGlyLeuGlnAspAArgPylsGlnValAArgGlnGly 501
513CGAGATTTCAACGCGCGCTGTGATGAGCAAGCTTGGACG 556
||:|||||:::|||||
502 LysGlyAArgGlyPheProLeuAArgGlyAArgGlyAArgAspAsnLeuGln 518
557 AAGCAAAATTCATGCTTGAAGCAAGCTGGCGCACAGCTGCCGTGTA 606
518 nHisAArgAArg.....CysLeuSerLeuAArgLeuGlnAArgProLysGlyAArgG 534
607 AATGCTGCCAATCATGCAACACATGAATTCGGCGCGCGCGCATCTCCGCG 656
:::|||||:::|||||
534 LysCysAArgGlyAArgGlnLeu.....LeuProGlnAArgAlaLeuAArg 549
657 TTTGAGTGGCACGACATTCATTCATGAGCC..... 689
550 ArgAsnProGlnGlyGlnAlaGlnHisAArgValAArgAArgSerLeuAArg 566
690GTCGCGCGCAATGAACACGCTGGACCAAT 723
566 nSerSerAlaSerGlnGlnAArgLeuAArgAArgAArgLeuHisGlnA 583
724 TATCAAGATGTAATTAC.....CATGCGCGTGT 755
||:|||||:::|||||
583 rGSerAArgGlnValPheGlyAlaGlnValProGlnHisThrProAArg 599
756 TCGAACAGCGCGTGAACACCGCGCGTGAATTCCTAGTGGTCTC 805
600 LeuProGlySerSerGlnLeuAArgLeuAArg..... 609
806 AAGTCAACAAACCGCGCTTTCGCTACCGTTTGGGTGCAAGTATCG 855
609 609
856 CAATATTACGCGCGGGAATGTGTGACACAGCAACCGCGTATTCGCG 905
610LeuGlnAlaAArgGlnPhe... 615
906 TTCGATATTGAACGCGCGATTCACAAAGCGCGCACGATATTGGGAC 955
:::|||||:::|||||
616PheAArgAArgGlySalAArgLysAArgHisAArgLeu...GlyAs 629
956 GCTACCAATCATGATTTCCGTTATCGAAGCAAGCGCGCGCAAGAGAGCTG 1005
||:|||||:::|||||
629 pLeuSerGlyGlyAspLeuAArgLeuAArgPylsAArgGlyGlnGln.....T 644

```
1006 TTGGCTGGGTTGGCCGCGAGACCCGAGCAAAATCTCCATTCACGGGTACAC 1055
      |||
644 LeuArg.....Gly 646
1056 CCTGGCCATTCTCTGAAAAAACAACCTTCACATTCACACAGCCGCTCA 1105
      |||
647 Pro.....GlyLeuArgHisArgLysArg 654
1106 ACGGGGCGAGCCGCGCCATGGTCCGATGTGTACTTACGACGCGGTGATG 1155
      |||
654 GlnSerArgPheAsnHis.....Pro 662
1156 CCCTGGATATCTGCCCCACCCCTGCTTTGGCGGATTATATCGTCGCCCA 1205
      |||
662 LaaIaValLeuProArgGluProCysSerGluAspGlyValArgGly 678
1206 TACCGACAGCGCGGCGCATTTGGGTTGCTTGAATTGACGAGA..... 1250
      |||
679 GlyLeuArgArgGlnGlyLeuProLeuGlyAlaAspArgGlnLeuArg 695
1251 .....AGACCTGCTTTGTGTGACGTTGTGTCGCCGCGCAATACGAAATC 1296
      |||
695 GValGlyProArgPhe.....GlnAspGluVal 705
1297 GGGCCGCTGTTGGCGAAAGTCTGGAACCATGAGAAAGA 1337
      |||
705 rProLeu.....SerGlySerHisAsnGlnGly 714
```

seq_name: /cgn2_6/ptodata/1/laa/6B_COMB.dep:US-09-413-814-92

seq_documentation_block:

```
; Sequence 92, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
; TITLE OF INVENTION: heteropolypeptide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; EARLIER FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-413-814-92
```

alignment_scores:

```
Quality: 94.00 Length: 536
Ratio: 0.423 Gaps: 30
Percent Similarity: 41.418 Percent Identity: 22.015
```

alignment_block:

US-09-303-518d-125 x US-09-413-814-92 ..

Align seq 1/1 to: US-09-413-814-92 from: 1 to: 608

28 CTGCCATCGGGGAGACCGGAGCAAGCGTTTACGACGCGCCGCCAT 77

```
1136 LeuGlnLeuSerHisProAspLeuAlaIleThrValAspProHisAl 152
      |||
78 TACCGAATCGCGCTGTTGGCGAAGAATAGCCGGATCGCCGCTCGA 127
      |||
152 AlaMetLeuAla.....SerGluysAspArgLeuAlaIleThrValSer 168
128 TGAAGTCAAGAGAGGCGATCCGTCAAAAAGCCCAAGTGTCTTGA 177
      |||
168 TyrLeu..... 170
178 GACAAAAAGATCCGCGCTGTGTATTACGCGCGCTTACGCGCAAT 227
      |||
171 AspLeuHisGlyProSerValValGlnThrAlaCysSerThrSer 187
128 CGCGCGCATTCACCGTGGCGAAGCGCGTACTTCACTGACGTCGTC...A 274
      |||
187 uValAlaValHisMetAlaCysGlnSerLeuGlnGlyGlnCysAsp 204
275 TTGCGCTTGAAGCGCAAGCGAATGAGTTGAACGC..... 312
      |||
204 euAlaLeuAlaGlyGlyValSerIleGlyIleProGlnLysArgGly 220
313 ...TACGACCTGAAAGCGCTGGCAACTTAAAGCGCGAA..... 348
      |||
221 ProTyrValProGlySerIleCysSerProAspArgLysArgProPh 237
349 .GAAGTGGCGCGCAACCTGATCCAAATCCGCT..... 378
      |||
237 eaAspAlaArgAlaGlnGlyThrValGlyGlySerGlyValGlyLeuVal 254
379 .....TTGTGACTGCGCTGCGCGACCGCTGCTG..... 408
      |||
254 laLeuysArgLeuAlaAspAlaLeuArgAspArgAsnThrValHisAla 270
409 .....ASCAAAAT..... 417
272 ValIleArgGlySerAlaValAsnAspArgLysArgLysValGlyPh 287
418 .....CCTGCGTGAATGCGGAGCCGCTGCGCATC..... 447
      |||
287 emetAlaProSerValAspArgGlnAlaAlaIleSerGlnAlaIle 304
448 .....TTGCTGCAATGCGATGAC 465
304 euValAlaGlyValAspProGlySerIleGlyLysValGlnAlaIle 320
466 ASCAAATCGCGTGGCTGCGCGACCTAGCGTATTCATGAAGAGCGCGCGA 515
      |||
321 ThrAlaThrAlaIleGlyAspPro...IleGlnValGlnAlaLeuThr 336
516 GGATTCAAAGCGCGCGCTGTGTATGAGCGCTTGAACGCAAGCAAA 565
      |||
336 naIaPheArg.....ArgLysThrProArgLys 346
566 TGCATGTTTGAAGCAAGTGGCGGACAGCGTGGCGGTAAGTGGTGGC 615
      |||
346 laTyr...Cys...AlaLeuGlySer.....IleLysAla 355
616 AACATCGAAACATGAAATGCGCGCGCGCATCTGCGGTTTGAAGTG 665
      |||
356 AsnIle...GlyHisLeuAspAlaAlaIleGlyValAlaIleGlyLeuIleLys 372
666 CACGCAATTCATTCATCGACGCGGCGGAGTAATAACCGTGT... 712
      |||
372 laAlaHisValVal.....ArgSerGlyGlnIleProCysVal 385
713 .....GACCAATCA..... 721
      |||
386 HisPheGlnAlaProAsnProLysLeuAspLeuAlaIleSerProPhe 402
722 .ATTATCAAGATGTAAATTCATTTGGCGG..... 749
      |||
```

```

402 eValProArGluAlaAlaPro.TPrProArGluLeuArGProArGArg 418
750 .....TTTGTTCACACAGCCGCTCTGACACCGAGCGG 784
419 AlaGlyValSerSerPheGlyIleGlyIleThrSerAlaHisValIle 435
785 TGATTCCCTAGGTGGTCTCAAGTCAACAAACCGCGCT..... 824
435 uGluGluProProLeuProLeuProArGAlaProAlaProGluArGAspH 452
825 .....CTTGGCTACCGCTTTGGTGGCGAAAGTATGGCAATTAC 863.
452 lValLeuThrLeuSerAlaArGThrPro...GluAlaLeuSerThrAla 467
864 TGGGGCGAATTGGTTCACACAGACACCGCGCTGAT..... 899
468 CysAlaGluLeuAlaHisLeuGluAlaThrAspValProLeuAspAs 484
900 .....TTCCGGTTCGGTATTGACCGCGCATTCACACAGCGCGCAC 942
484 pValAlaPheThrLeuGluIleThrGlyArGAlaHisProGlyArGAsp 501
943 GATTATTGGAGCGCTACCAATTCAGATTTCGGTATCGAAGAGCGG 992
501 lAlaValAla.....ArgThrArGAlaGluAla 510
993 CAGCAAGACGCTGTGCGTGGGTGGCGCGACCGCGCAAAATCTCCA 1042
511 lIleGluIleLeuAlaArGAlaArGAlaSerAla.....LeuAl 523
1043 TCACGCGTACACCCCTGGCCATTCTCTGAAAAACAACCTCTCAAGTTC 1092
523 aArGProAspGluProArGPro..... 530
1093 AACACACGCGTCAACGCGCGACCGCGCATGTGCGCATGTGACTTA 1142
531 ..SerSerArGSerAlaArGAlaArGArGProSerGlyTTP..... 544
1143 CGAGCGCGTGTATGCCCTGGATATCTGCCACCCCTGCTTTCCGGGATT 1192
545 .....Pro.....AlaArGSe 548
1193 TAATCGTCGGGATACGACACGCGAGCATGTGGTCTTGGAATTG 1242
548 TThrTrArGArG.....ArgSerGlyAlaProSerThrArGAla 562
1243 GACGAGAAGACCTCGCTT.....TGTGCACTTGCTGCC 1279
563 rGArGArGArGlyArGAlaAlaSerIleSerAlaArGSerSerAla 578
1280 CGGCGCAATACGAATACGCGCGCTGTGGCAAAAGTCTGGAACCATT 1329
579 ArGAlaAlaArGAlaThrGlyAlaArGysSerAlaProArGTrpArGSerP 595
1330 GAGA 1333
595 oArG 596

```

```

seq_name: /cgn2_6/ptodata/1/iaa/5b_COMB.pep.us-08-878-957-32
seq_documentation_block:
: Sequence 32, Application US/08878957
: Patent No. 5965796

```

GENERAL INFORMATION:

```

APPLICANT: Meagher, Richard B.
APPLICANT: Summers, Anne O.
TITLE OF INVENTION: Metal Resistance Sequences and
TITLE OF INVENTION: Transgenic Plants
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201

```

```

CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,957
FILING DATE: 19-JUN-1997
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/427,097
FILING DATE: 21-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Feider, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 40-94A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
TELEX:
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-957-32

```

```

alignment_scores:
Quality: 93.00 Length: 421
Ratio: 0.484 Gaps: 17
Percent Similarity: 45.606 Percent Identity: 20.190

```

alignment_block:

```
US-09-303-518D-125/rev x US-08-878-957-32 ..
```

```
Align seg 1/1 to: US-08-878-957-32 from: 1 to: 564
```

```

1122 GCGCGCGTCCGCCGCTGACGCGCTGTGGAACCTGAAGAGTTGTTT 1073
113 GAGGCTGTTGTTT.....GAGGCTGTACGCGTGTGAG 1038
1072 TCAGGAATGAGC.....GAGGCTGTACGCGTGTGAG 1038
179 aGlnGlnAlaAlaArGValAspGluLeuArGHisAlaLysTYrGluGlyI 196
1037 TATTGTCCGCTCGCGCGACACCCGACCGACGAGTCTT.....GCT 994
196 le.....LeuGluGlyAsnProAlaIleThrValIleuHisGlySer 209
993 GCGGCTCTCTGATACGGAATCTGATGTGTA.....GC 956
210 AlaArGpHeuYsAspAsnArGAsnLeuIleValGlnLeuAsnAspGlyI 226
955 GTCCCAATTAATCGTCCGCGCTGTGTAATCGCGC..... 919
226 yGluArGValAlaAlaPheAspArGysLeuIleAlaThrGlyAlaSerP 243
918 .....GTTCAATACGGAACCGGA 901
243 roAlaValProProIleProGlyLeuYsAspThrProTYrTTPThrSer 259
900 AATCAAGCGGTGTCTGT..... 883
260 ThrGluAlaLeuValSerGluThrIleProLysArGLeuAlaValIleG 276
883 ..... 883

```


428 ATGCCGAGCCGTTGCCATCTTGTCAATGCGATGAGCAACCAATCCGCTG 477
 815 rgtatargproglvalvalvalaserasergerglvalval 831
 478 GCTGCC.....GACCTTACGTCATTAATCAAGACCGCCGA 515
 832 glyserglvalleuglvalaspprogltytrtrvalarghisalargl 848
 516 GGAATTCAAA.....CGCGGCTGTGGTAT 541
 848 valavalargprhealaasglvalglvalvalarglvalleuglval 864
 542 TGAGCCGTTTGACGACGCAAAATCCAT..... 570
 865 ..gltphleuvalglvalglvalprohisglvalleuthglmetala 880
 571 ..GTTTGAAGCAGCTGGCGCAGAC.....GTCCGCTCGAAA 608
 881 glglucylleuglvalglvalasprvalvalvalprolalemtar 897
 609 TGCTGCAACATGCAACACATGATTCGGCGC..... 642
 897 gtagglvalglvalglvalargglvalphelualalaleualathvalp 914
 643 ..CCGATCCTCCGCGTTTGAAGTGCACGACATTCATTCAGCGCG 690
 914 hethrargaspralaglvalleualatthrthleunis...ThglYser 929
 691 GTCCGCGCAATGAAACCGTGGACCATTAATCAAGATTAATTC 740
 930 ThglYarglglleaspleuprothrtprprohegnhis..... 943
 741 CATTTGCGCT.....TTGTTGCAACGCGCTGTGACACCGCGCG 784
 944 ..Aspargttrtrleualalaproserarproartrarglval 959
 785 TGATGGCCGAGGTGTTCAATGCAACGCGCTGTGCGTAC 834
 959 euseralalaglvalleualarglvalglhisproleuethrvalala 975
 835 GTTTGGGTGCAAGATGCAATTAATTCGCGCGCAATTCGTTGACAC 884
 976 Val.....GluleuproglYth 981
 885 AGAACAACCGGTATTCGGTTCGGATTAAGACGCGCATTAACAAG 934
 981 rAspThrGlualThrThglYarglleseralalalasprouprotrp 998
 935 GCGCG...CAAGATTATTTGGACGCTACCAATC..... 967
 998 eualalaspHis..leuValtrpararglvalvalproglYthvalale 1014
 968AGATTTCGTTATCAAGACG 989
 1014 uleucluthvalleuclnvalglYserarglglYleuproargvala 1031
 990 CC..... 991
 1031 lagluleuvalleucluthrproleuthrtrparaserparproleu 1047
 992GCGCAAAAGCTGTGGCTGGGTGCG 1020
 1048 glnvalarglvalvalthrvalalalathrvalproglYglYala 1064
 1021 CCGGACCGGACAAATCTCCATCAGCGTACACCTCGGCAATTTCT 1070
 1064 ggluleuthrleunis.....SerargprogluProvalala..... 1076
 1071 GAAAAACAACCTTCAGATTCAACACGCGGTACAGCGCGCGCG 1120
 1077AlaserSerSerProserProalaser.Proar 1088

1121 CCAT.....GTCGCGATTGTACTTACGCGCGCTG 1152
 1088 ghisLeuthAlagIngluseraspraprrprrharghisalaserG 1105
 1153 ATGCCCTTGATATCTGCC..... 1172
 1105 lyleuLeualalProalalaglYleualalaspasprhealagluleuthr 1121
 1173CA 1174
 1122 glvalatrrprovalglYalaglproleualasphealaglYalnty 1138
 1175 CCCGCTTTTGGCGGATTAATCGTCGCGATACCGACGCGCGCGCA 1224
 1138 rProleuphehialalaglYalarg.....TyrGluglYalalphe 1153
 1225 TTGGTTCCTTGGATTGACGAAAGACCGTGC...TTTGTGACGCT 1271
 1153 rglYleualarglalaatrrpararglYasprglvalalphealaspval 1169
 1272 CGTCGCGC 1280
 1170 ArgLeupro 1172
 seq_name: /cgn2_6/prodata/1/iaa/5B_COMB.pep:us-08-804-198-3
 seq_documentation_block:
 ; Sequence 3, Application us/08804198
 ; Patent No. 5945320
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgett, Stanley G.
 ; APPLICANT: Kuhstoss, Stuart A.
 ; APPLICANT: Rao, Nagaraja R.
 ; APPLICANT: Richardson, Mark A.
 ; APPLICANT: Rostock, Paul R., Jr.
 ; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PAUL R. CANTRELL 1138
 ; STREET: LILLY CORPORATE CENTER
 ; CITY: INDIANAPOLIS
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: Macintosh 7.0
 ; SOFTWARE: Microsoft Word 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,198
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CANTRELL, PAUL R.
 ; REGISTRATION NUMBER: 36,470
 ; REFERENCE/DOCKET NUMBER: P9113
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-276-3885
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1996 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-804-198-3
 alignment_scores: 471
 Quality: 93.00
 Ratio: 0.472
 Percent Similarity: 41.826
 Length: 23
 Gaps: 23
 Percent Identity: 22.081

alignment_block:

US-09-303-518D-125 x US-08-804-198-3

Align seq 1/1 to: US-08-804-198-3 from: 1 to: 1996

```

199 GTGTTTCTGCGCGGCTTCAGCAAAATGCGCGGATTCACCGTGGCGA 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
749 ValSerValAlaAlaValAlaSerGlyProArgSerValValLeuSerGlyAs 765
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 AAGCGCGTACTTCAGTCAGTGGTTCATTCGCGTGAAGCAACGAGAA 298
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
765 pValGlyValLeuGluSerValAlaAlaSerLeuMetGlyAsp..... 779
299 TCAGATTTCAGCGCTACGACCTGAAAGCGCGTGAAGCAATTAGCGCGCA 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
780 .....GlyVal 781
349 GAATGGCGCGCACTGATCCATCCGCTTTGTGACTGCGCGTGGCGAC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
782 GluCyArgArgLeuAspValSerHisGlyPheHisSerValLeuMetGly 798
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
399 CGGTCCGTTACGAAATTCCTGCCGTC.....G 427
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
798 uProValLeuGlyGluPheArgGlyValValGluSerLeuGluPheGly 815
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428 ATGCCGAGCGGCTTCGCTTCATTCGATGCGAGACCAACCGCGCTG 477
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
815 rGValArgProGlyValValValSerSerValSerGlyValVal 831
478 GGTGCC.....GACCTACGCTCATTCACAAAGACGCGCGA 515
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
832 GlySerGlyGluLeuGlyAspProGlyTyrTrpValArgHisAlaArg 848
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
516 GAATTTCAA.....CGCGCGCTGTGGAT 541
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
848 uAlaValArgPheAlaAspGlyValGlyValValArgGlyLeuGlyVal 864
542 TGAGCGCTTGACCGCAACGCAATTCAT..... 570
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
865 ..GlyThrLeuValGlyValGlyProHisGlyValLeuThrGlyMetAla 880
571 ...GTTGTAAAGCGAGTGGCGCGAGC.....GTGCGCTGTGAAA 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
881 GlyGluCysLeuGlyAlaGlyAspAspValValValProAlaMetArg 897
609 TGCTGCCACATCGAAACACATGATTCGGCGGC..... 642
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
897 gArgGlyArgAlaGluArgGluValPheGluAlaAlaLeuAlaThrValP 914
643 ..CGCATCTTCGCGGTTGTGAGTGCACGACATTCATTTCATCGAGCG 690
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
914 heThrArgAspAlaGlyLeuAspAlaThrThrLeuHis...ThrGlySer 929
691 GTGCGCGCGAATAAACCGTGTGACCATCAATTCATCATCATTAATAC 740
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
930 ThrGlyArgArgLeuAspLeuProThrThrProPheGluHis..... 943
741 CATGGCGGT.....TTGTTGCAACAGCGCTGTGAACCGCGAGCGG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
944 ...AspArgTyrTrpLeuAlaAlaProSerArgProArgThrAspGlyL 959
785 TGATTGCCCTAGGTGTTCAGTCACAAACACCGCGCTTCGTGGTAC 834
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
959 eSerAlaAlaGlyLeuArgGlyValGluHisProLeuLeuThrAlaAla 975
835 GTTTGGGTGCGAAAGTATCGCAAAATTCATGCGCGGCAATTGTTGAC 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
976 Val.....GluLeuProGlyThr 981
885 AGACAACCGCGTATTCGCTTCATTCGAGCGCGGATTCACAAAG 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
981 rAspThrGluValTrpThrGlyArgGlySerAlaAlaAspLeuProTyrL

```

```

935 GCGCG...CACGATTATTGGACGCTACCAACATC..... 967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
998 euAlaAspHis..LeuValTrpAspArgGlyValValProGlyThrAlaLe 1014
968 .....AGATTCCGTTATTCAGGAAG 989
1014 uLeuGluThrValLeuGluValGlySerArgGlyLeuProArgValAla 1031
990 CC..... 991
1031 laGluLeuValLeuGluThrProLeuThrTrpThrSerAspArgProLeu 1047
992 .....GCAGCAAGAGCTGTGCGTGGCTGG 1020
1048 GluValArgGlyValValThrAlaAlaAlaThrAlaProGlyGlyAlaArg 1064
1021 CCGCAGCGCGCAAAATTCATTCAGCGGTACACCGCTGCGCATTTGCT 1070
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1064 gGluLeuThrLeuHis.....SerArgProGluProValAla..... 1076
1071 GAAAAACAACCTCTCAAGTTCACACAGCGCGTACAGCGCGCGCGG 1120
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1077 .....AlaSerSerSerSerProSerProAlaSer..ProArg 1088
1121 CCAT.....GTTGCCGATTGCTTACGACGCGGTG 1152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1088 gHisLeuThrAlaGluGluSerAspAspAspTrpThrArgHisAlaSerG 1105
1153 ATGCCCTTGATATCTGCG..... 1172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1105 lYLeuLeuAlaProAlaAlaGlyLeuAlaAspAspPheAlaGluLeuThr 1121
1173 .....CA 1174
1122 GluAlaTrpProValAlaGlyAlaGluProLeuAspLeuAlaGlyGly 1138
1175 CCCGCTTTGGCGGATTTATCTCGCGCATTCGACGCGCGAGCGA 1224
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1138 rProLeuPheAlaAlaGlyValArg.....TyrGluGlyAlaPheAla 1153
1225 TTGGTTCCTTGAATTGACGAGAGACCTGCG...TTGTGCGAGCTT 1271
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1153 rGlyLeuArgAlaAlaAlaTrpArgArgGlyAspGlyValPheAlaAspVal 1169
1272 GGTCTGCC 1280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1170 ArgLeuPro 1172
seq_name: /cgn2_6/prodata/1/1aa/5B_COMB.pep:US-08-467-822-26
seq_documentation_block:
; Sequence 26, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Rentero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,822
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4000
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 569 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-822-26

```

```

alignment_scores:
Quality: 92.50 Length: 424
Ratio: 0.517 Gaps: 22
Percent Similarity: 42.217 Percent Identity: 20.519

```

```

alignment_block:
US-09-303-518D-125 x US-08-467-822-26 ..

```

```

Align seg 1/1 to: US-08-467-822-26 from: 1 to: 569

```

```

205 ACTGCGCGCGCTGACGCAAAATCGCGCATTCACCGTGGCGAAAGCG 254
|||||
161 ThrGlyProAlaAspGlyThrAsnAlaThrThrIleThrProGlyArg 177
225 CGTACTTCAGTCAGTCGATGCGGATTCGCGTGAAGCAACGACGAATGACG 304
| |||||
177 gasnleuylstrpmetleuargalala 186
305 TTGAACGCTACGACCTGAA.....GGCTGGCAAACTTAAGC 342
|||||
167 .GlucIlyrSerMetAsnleuGlyPheLeuAlaIalysGlyAsnAlaSer 202
343 GCGGAGAAGTGGCGCGACCTGATCCAAATCCGCTTG..... 361
|||||
203 AsnAspAlaSerAlaArgAspGlnIleGluAlaGlyAlaIleGlyPhe 219
382 .....TGACTGCGCTGGCGACCCGTCGCTGACGCAAAATTC 418
|||
219 sIleHisIleuAspTrp.....GlyThrThrProSerAlaIleAsn 233
419 CTGCGCGTGCAT...GCGAGCGGCTTCGCGCATCTTGCAATGCGATGAC 465
|||||
233 IAlaIleuAspValAlaAspIlyAspValGlnVal...AlaIleHis 248
466 ACCAATCGCGTGGCTGGCGACCGCTACGTCATTCGAAAGAGCC..... 510
|||||
249 ThrAspThrLeu.....AsnGluAlaGlyCys 257
511 .GCGGAGATTTCAAACGCGCGCTGTGTATTGAGCGTTGACCGAAC 559
|||||
257 sValGluAsp.....ThrMetAlaIalIleAlaGly 268

```

```

560 GCAAAATCCATCTTTGTAAGCGAGTGGCGACGACGTGCGCTGAAAT 609
|||
268 rGthrMetHisThrPheHisThrGluGlyAlaGly..... 279
610 GCTGCCAACATCGAACAACATGAAATCGCGCGCGCATCTCGCGTTT 659
|||||
280 .....GlyGlyHisAlaProAspIle 287
660 GAGTGGCAGCGACATTCATTCATCGAGCGGTGGCGCGCAATTAACCG 709
|||||
287 elysValAlaGlyGluHisAsnIleLeuProAlaSerThrAsnProThr 304
710 TG...TGACCATCAAT.....TATCAAGATGTAATTAACATT 744
|||||
304 IeropherThrValAsnThrGluAlaGluHisMetAspMetLeuMetVal 320
745 GCGCGT.....TTGTTTGGACACAG 764
321 CysHisHisLeuAspIlySerIleGlyGluAspValGlnPheAlaAsp 337
765 CGGTTCGACACCGAGCGCGTGTATGCC.....CTAG 796
|||||
337 rArgIleArgProGlnThrIleAlaAlaGluAspThrLeuHisAspMet 354
797 GTGTTTCGAGTCAACACAAACCGCG..... 822
|||
354 IylPheSerIleThrSerAspSerGlnAlaMetGlyArgValGly 370
823 .....CTCTGGCTACCGTTTGGTGGCAAGATGCAAAATTAATCTGC 866
371 GluValIleThrArgThrTrpGlnThrAlaAspIlyAsnIlyGluPh 387
867 GGGCGAATTGGTT.....GACACAGACAAAC.....C 892
|||||
387 eGlyArgLeuIlyGluGluIlyGlyAspAsnAspPheArgIlyLys 404
893 GCGTGAATTCGCGTTCGATGACGCGCGATTCACACAGCGCGCAC 942
|||
404 rGlyrLeuSerIlyThrIleAsnProAlaIleAlaHisIlyIleSer 420
943 GATTATTTGGAGCTACACCAATCAATTCGTTATGCAAGAGCGCG 992
|||||
421 GluTrpValGlySer.....ValGluValGly 430
993 CAGCAAGAGCTGTGCGCTGG..... 1014
430 sValAlaAspLeuValLeuTrpSerProAlaPheGlyValLysPro 447
1014 ..... 1014
447 smetIleIleIlyGlyGlyPheIleAlaLeuSerGlnMetGlyAspAla 463
1015 .....GTTGCGCGCGACGCGGACAAATACTCCATCAGCGGTAC 1052
|||||
464 AsnAlaSerIleProThrProGlnProValIlyTrp.....Arg 477
1053 AACCTCGGCGCAT.....TTCGGAAGAAACAACTCTTCA 1087
|||||
477 umetPheGlyHisHisIlyLysAlaIlyLysIlyAspAlaGlnIleThrPhe 494
1088 AGTTCAACACACCGCTCAACGCGCGCGCGCCATGTCGCGGATGGT 1137
|||||
494 alserGlnAlaIalAlaIlyAspIlyGlyIleGlyGlu.....GluLeuGly 508
1138 ACTTACGACCGCGTGAATGCGCTTGATATCTGCGCACCGCTGTTGG 1187
|||||
509 leuGluArgGlnValLeuProValLysAsnGlyAsnGlnIleThrIly 525
1188 CGATTAACTCGCGCGATACC 1209
525 sasPheGlnPheAsnAspThr 532

```

seq_name: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:us-08-432-697-26

seq_documentation_block:

Sequence 26, Application US/08432697
Patent No. 6248330

GENERAL INFORMATION:

APPLICANT: Ladique, Agnes

APPLICANT: Sauerdam, Sebastien

APPLICANT: Ferrero, Richard L.

APPLICANT: Thibierge, Jean-Michel

TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST

TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE

TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/432,697

FILING DATE: 02-MAY-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 03495.0137-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4400

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 569 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-432-697-26

Alignment scores:

Quality: 92.50

Ratio: 0.517

Percent Similarity: 42.217

Percent Identity: 20.519

alignment_block:

US-09-303-518D-125 x US-08-432-697-26

Align seq 1/1 to: US-08-432-697-26 from: 1 to: 569

```

205  ACTGGCGCGGCTTCAGCAAAATGCGCGCATTCACCGTGGCAAAAGCG 254
161  ThrGlyProAlaAspGlyThrAsnAlaThrThrProGlyArgat 177
255  CGTACTTCAGTACGTGATTCGCGTTCGGAAGCAACGACGAATGAGT 304
177  GAsnLeuIysTrpMetLeuArgAlaAla 186
305  TTGAACGCTGCGACCTGAA 342
187  ..GlugIuTySerMetAsnLeuGlyPheLeuAlaLysGlyAsnAlaSer 202
343  GGCAGAAAGTGGCGCGCATGATCCAAATCGGTTTG 381

```

```

203  AsnAspAlaSerAlaArgAspGlnIleGluAlaGlyAlaIleGlyPhe 219
382  .....TGACATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418
219  sIleHisGluAspTrp.....GlyThrThrProSerAlaIleAsn 233
419  CTGGCGCGCGAT...GCCGACCGCTGCGCGCGCGCGCGCGCGCGCGCG 465
233  IsAlaLeuAspValAlaAspLysTyArgValGlnVal...AlaIleHis 248
466  ACCAATCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 510
249  ThrSprThrLeu.....AsnGluAlaGly 257
511  .GCCGAGATTTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 559
257  sValGluAsp.....ThreAlaAlaIleAlaGly 268
560  GCAAAATTCATGTTTGTAAAGCAGCTGGCGCGCGCGCGCGCGCGCG 609
268  rGhrMetHisThrPheHisThrGluGlyAlaGly..... 279
610  GTGCCCAATCGAACAACATGATTCGCGCGCGCGCGCGCGCGCGCG 659
280  .....GlyGlnHisAlaProAspIleI 287
660  GAGTGGCAGCAGCATTCATTCATGAGCGCGCGCGCGCGCGCGCGCG 709
287  eLysValAlaGlyGlnHisAsnIleLeuProAlaSerThrAsnProI 304
710  TG...TGACATCAAT.....TATCAAGATGATATACAT 744
304  LeProPheThrValAsnThrGluAlaGlnHisMetAspMetLeuVal 320
745  GGCGCG.....TGTTGGCAACAG 764
321  CysHisHisLeuAspLysSerIleLysGluAspValGlnPheAlaAsp 337
765  CGGTGGAACACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 379
337  rArgLeuArgProGlnThrIleAlaAlaGluAspThrLeuHisAsp 354
797  GTGGTTTCAGAGTCAACAACCGCGC 822
354  LylIlePheSerIleThrSerSerAspSerGlnAlaMetGlyArgVal 370
823  ....CTCTGCGGTACCGTTTGGGTGCGAAGTATCGCAAAATTACTGC 866
371  GluValIleThrArgThrTrpGlnThrAlaAspLysAsnLysGluP 387
867  GGGCGAATTGGTT.....GACACAGACAC 892
387  eGlyArgLeuLysGluGluLysGlyAspAsnAspAsnPheArgIle 404
893  GCGTGATTTCCGCTTCGATATTGACGCGCGCGCGCGCGCGCGCGCG 942
404  rGlyrLeuSerLysTyThrIleAsnProAlaIleAlaHisGlyIle 420
943  GATATTTGGGACCGCTACCAATCAGATTTCGGTATTCGCAAGAGCGCG 992
421  GluTyValGlySer.....ValGluValGly 430
993  CAGCAAGAGCGCTTCGCGTGG..... 1014
430  sValAlaAspLeuValLeuThrProAlaPhePheGlyValLysPro 447
1014  ..... 1014
447  snMetIleIleLysGlyGlyPheIleAlaLeuSerGlnMetGlyAsp 463
1015  .....GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1052
464  AsnAlaSerIleProThrProGlnProValTyTyTyr.....ArgI 477

```

1053 AACCTCGGCAT.....TTCGTAAACAACTCTCA 1087
 477 unethrhegylhshisgllyslalylslyrprgslrphnlethrrphev 494
 1088 AGTTACACACGCGCGTCAACGCGGCGACGCGCATGTGTCGATGTGT 1137
 494 alserglnlaalaatyrsplysgllylelysglu.....gluueuly 508
 1138 ACTTACGAGCGCGTGTGCGCTGTGATATCCGCCACCTGCTTTTGGC 1187
 509 leuglrlarglnvalleuprovalllysasnrcysargasnilethrllysl 525
 1188 CGATTTAATCGTCGCGCATACC 1209
 525 sasmetglnpheasnspthr 532

seq_name: /cgn2_6/ptodata/1/laa/6b_COMB.pep:us-08-466-248-26

seq_documentation_block:

; Sequence 26, Application US/08466248

; Patent No. 6258359

; GENERAL INFORMATION:

APPLICANT: Labigne, Agnes

APPLICANT: Sauerbaum, Sebastien

APPLICANT: Ferrero, Richard L.

APPLICANT: Thiberge, Jean-Michel

TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST

TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE

TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/08/466,248

APPLICATION NUMBER: US/08/466,248

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 08/447,177

FILING DATE: 19-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/432,697

FILING DATE: 02-MAY-1995

CLASSIFICATION: 435

TELECOMMUNICATION INFORMATION:

REFERENCE/DOCKET NUMBER: 03495.0137-02000

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 569 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

alignment_scores:

Quality: 92.50 Length: 424

Ratio: 0.517 Gaps: 22

Percent Similarity: 42.217 Percent Identity: 20.519

alignment_block:

US-09-303-518d-125 x US-08-466-248-26

Align seg 1/1 to: US-08-466-248-26 from: 1 to: 569

205 ACTGCGCGCGCTTCAAGCAAAATGCGCGCATTCACCGTCGCGCAAAAGC 254
 161 ThrGlyProAlaAspGlyThrAsnAlaThrThrIleThrProGlyArg 177
 255 CGTACTGTCAGTCGTCGATTCGCGTCAAGCAAGCAAGCAAAATCGAGT 304
 177 gsnleuylstrpmetleuargalala..... 186
 305 TTGAACGTACGACCTGAA.....CGCTGGCAAACTTAAAC 342
 187 ..gluGluTyrSerMetAsnLeuGlyPheLeuAlaLysGlyAsnAlaSer 202
 343 GCGGAAGAAGTGGCGCGCAACCTGATCCAAATCCGGTTG..... 381
 203 AsnAspAlaSerAlaArgAspGlnIleGluAlaGlyAlaIleGlyPhe 219
 382TGACTGCGCTGCGCACCGCGCTTCAGCAAAATTC 418
 219 slleHsIgluAspTrp.....GlyThrThrProSerAlaIleAsn 233
 419 CTGCGCTGCAT..GCCGACCGCTTGGCATCTGTCGAAGCGATGAGC 465
 233 IsAlaLeuAspValAlaAspLysTyrAspValGlnVal...AlaIleHs 248
 466 ACCAATCCGCTGCGCTGCGCACCTTACGTCATTATCAAGAAGCC..... 510
 249 ThrAspThrLeu.....AsnGluAlaGlyCys 257
 511 ..GCCGAGATTTCAAACGGCGCTGTGATTCAGCCGTTCAGCGAAC 559
 257 svalGluAsp.....ThrMetAlaAlaIleAlaGly 268
 560 GCAAAATCCATTTTGTGAAGCAGTGGCGCAAGCTGCGCTGTAAT 609
 268 rglThrMetIleThrPheHsIleThrGluGlyAlaGly..... 279
 610 GCTGCAACATCGAAACACATGAATTCGGCGCGCGCGCATCTGCCGTT 659
 280GlyGlyHsAlaProAspIleI 287
 660 GAGTGGCAGCGACATTTATTCATCGACCGCGCGCGGATTAACCG 709
 287 elysValAlaGlyGluHsAsnIleLeuProAlaSerThrAspProThrI 304
 710 TG...TGACCATCAAT.....TATCAAGATGTAATTCACAT 744
 304 leProPheThrValAsnThrGluAlaGluHsMetAspMetLeuMetVal 320
 745 GCGCGT.....TTGTTTGCACACAG 764
 321 CysHsIleAspLysSerIleLysGluAspValGlnPheAlaAsp 337
 765 CCGTGTGAACACGACGCGCGTGAATTC.....CTAG 796
 337 rArgIleArgProGlnThrIleAlaAlaGluAspThrLeuHsAspMet 354
 797 GTGTTTCAGTCACAAACCGCGC..... 822
 354 lylIlePheSerIleThrSerSerAspSerGlnAlaMetGlyArgValGly 370
 823CTTTGCGTACCGTTTGGTGGCGCAAGATACGCAATTTACTGC 866

```

371 GluValIleThrArgThrTrpGlnThrAlaAspLysAsnLysLysGlnPh 387
867 GGGGGAATTGGTT.....GACACAGACAAC.....C 892
387 eglValGluLeuLysGluGlnLysGlyAspAsnAspAsnPhenArgIleLysA 404
893 GCGTGATTTCCGGTTCGATTGACGGCGGATTCACACAGACGGCGGAC 942
404 rgtYrLeuSerLysTrpThrIleAsnProAlaIleAlaHisGlyLysSer 420
943 GATTATTGGAGACGCTACACAAATCAGATTTCGTTATCGAAGAGCGCG 992
421 GluTrpValGlySer.....ValGluValGlyLys 430
993 CAGCAAGAGCTGTCGCGCTG..... 1014
430 sValAlaAspLeuValLeuTrpSerProAlaPheGlyValLysProA 447
1014 ..... 1014
447 smetIleIleLysGlyPheIleAlaLeuSerGlnMetGlyAspAla 463
1015 .....GTGCGCGCGGACGGGACAAATCTCATACGCGTAC 1052
464 AsnAlaSerIleProThrProGlnProValTrpYr.....ArgG 477
1053 AACCTCGGCAT.....TTCGTGAAACAAACTCTCA 1087
477 umetPheGlyHisHisGlyLysAlaLysTrpAspArgAsnIleThrPheV 494
1088 AGTTCACACAGCCGCTCAACGGCGGACCGCCATGTCGCGATTTGCT 1137
494 alserGlnAlaAlaLysTrpLysGlyLysGlu.....GluLeuGly 508
1138 ACTTACGAGCGCGTATGCTCCCTGATATCTCTGCCACCGTGGTTGCG 1187
509 LeuGlnArgGlnValLeuProValLysAsnLysArgAsnIleThrLys 525
1188 CGATTATGTCGCGCGCATAC 1209
525 sAspMetGlnPheAsnAspThr 532

seq_name: /cgn2_6/prodata/1/1aa/B_B_COMB.pep:US-08-669-785-4
seq_documentation_block:
; Sequence 4, Application US/08669785
; Patent No. 6309668
; GENERAL INFORMATION:
; APPLICANT: Betsou, Fotini
; APPLICANT: Sebou, Peter
; TITLE OF INVENTION: Protective Epitopes of Adenyl
; TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
; TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections.
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.

```

```

; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1705 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-669-785-4

```

```

alignment_scores:
Quality: 92.50 Length: 467
Ratio: 0.467 Gaps: 24
Percent Similarity: 42.398 Percent Identity: 21.413

```

alignment_block:

```
US-09-303-518d-125 x US-08-669-785-4 ..
```

```
Align seg 1/1 to: US-08-669-785-4 from: 1 to: 1705
```

```

67 GCGCCGCCATTACCGAAGTCGCTGCTTGGCGAAGATATCGCGTAT 116
117 TCGTGTTCGAGACAAAGATCCGGC..... 195
801 a.....PheValAspArgPheIleGlnGlyLysArgValAlaGlyGlnPro 816
196 GTGGTGTTCGCGCGCGCTTCAGGCAAAATCGCGCATTCACCGTGG 245
817 ValValLeuAspValAlaAlaGlyGlyIleAspIleAlaSerArgLys 833
246 CGAAAGCGCGTACTTCAGTCAGTCGATTCGCGGTTGCGTGAAGGCAACGACG 295
833 yGluArgProAlaLeuThrPheIleThrProLeuAla..... 845
296 AAATCGAGTTGAACGCTACGACCTGAAGCGCTGGCAACTTAAGCGCG 345
846 .....AlaPro.....Gly 848
346 GAAGAGTCGCGCGCAACCTGATCAATCCGTTGTGGACTGCGCTGG 395
849 GluGluGlnArgYr.....Ar 854
396 CACCGCGCGTTCAGCAAAATTCCTGCGTCGATGCGGACCGCTTGCGCA 445
854 gTrpLys.....ThrGlySerGlnPheTrp 864
446 TCTTCGCAATGCGATGACACAAATCCGCTG..... 477
864 hrPheValGluIleValGlyLysGlnAspArgTrpArgIleArgAspGly 880
478 GCTGCGGACCTACGCTCATATCAAGAGAGCGCGGAGATTCACAAAG 527
881 AlaAlaAspThrThrIleAspLeuAlaLys..... 890
528 CGGCGTGTTCGATGAGCGCTTCACGGAAGC..... 561
891 .....ValValSerGlnLeuValAspAlaAsnGlyValLeuLys 904
562 .....AAATCGATGTTGTGAAGCGAGCTGCGGACGATCGCT 603
904 lserIleLysLeuGlnValIleGlyGlyAspGlyAspAspValValLeu 920

```

```

604 GAAATGCTGCCACATCGAACAACATGATTCGGCGCCGCTCTGC 653
    |||||
921 AlaasnlaserArgile.....Histyras 929
654 CGGTTGAGGGGACGACATTCATTTCAGCGCGGTGGCGGATA 703
    |||||
929 pgllyglialaglyThrAsn..... 935
704 AACCGTGTGGACCATCAATTAT.....CAMANGTAATT 738
    |||||
936 .....ThrValSerTyrAlaAlaLeuGlyArgGlnAspSerile 948
739 ACCATTGGCGGTTGTTTGCACAGGC...CGTGTGACACCGACGGGT 785
    |||||
949 ThrValSer.....AlaAspGlyGluArgPheAsnValArgLys... 961
786 GATTGCCCTAGTGTCTCAAGTCAACAACCGCCCTTTCGT... 831
    |||||
962 .....GlnLeuAsnAsnAlaAsnValTyrArgGluG 972
832 .....ACCGTTTGGTGGCAAGATACGCAATT... 861
972 LyvalAlaThrGlnLysThrAlaTyrGlyLysArgThrGlnAsnValGln 988
862 .....ACTGCGGCGAATTGTTGACACAGACA 890
989 TyrArgHisValGluLeuAlaArgValGlnLeuValGlnValAspH 1005
891 C.....CGCGTATTCGGTGGTATTCGAGCGCGGATTA 928
    |||||
1005 rLeuGlnHisValGlnHisIleIleGlyGlyAlaGlyAsnAspSerileT 1022
929 CACAGCGCGGCGAATTATTTG..... 951
1022 hrGlyAsnAlaHisAspAsnPheLeuAlaGlyLysAspAspArg 1038
952 .....GGAGCGTACACAAATCAGATTTCGTTATCGAAGACGCGC... 993
1039 LeuAspGlyGlyAlaGlyAsnAspThrLeuValGlyGluGlyHisAs 1055
993 ..... 993
1055 nThrValValGlyLysAspAspValPheLeuGlnAspLeuGly 1072
994 .....ACGAAGAGCGTGTGCGGTGCGGCGGACGCGCAAAATAC 1038
    |||||
1072 aITPpSerAsnGlnLeuAspGlyGlyAlaGlyAspThrValLysTyr 1088
1039 TCCATCAGCGGTACACACCGCTCGCCATTTCCTGAAAACAACTCTTCA 1088
    |||||
1089 AsnValHisGlnProSer.....GlnGluArgLeuGluArg 1100
1089 GTTCAACACAGCGGTCAACGCGGCGGACCGCGCATGTCGCGATGTA 1138
    |||||
1100 gmeGly.....AspThrGlyIleHisAlaAspLeuGlnLysGly 1114
1139 CTTCAGAGCGC.....GTGATG 1155
    |||||
1114 hrValGlnLysTrpProAlaLeuAsnLeuPheSerValAspHisValLys 1130
1156 CCGCTGAGTATCTGCGCCACCGCTTTCGCGATTTATCTGCGGCGA 1205
    |||||
1131 AsnIleGlnAsnLeuHisGlySerSerLeuAsnAspSerIleAlaGlyAs 1147
1206 T 1206
    |||||
1147 p 1147

```

seq_name: /cgr2_6/protdat1/1aa/3A.COMB.pep:US-08-476-008-60

seq_documentation_block:
; Sequence 60, Application US/08476008
; Patent No. 5627061

```

GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. B44F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,008
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6047
TELEFAX: (314)537-6099
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 427 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-60

alignment_scores:
Quality: 92.00 Length: 431
Ratio: 0.474 Gaps: 21
Percent Similarity: 45.012 Percent Identity: 20.882

Alignment_block:
US-09-303-518D-125 x US-08-476-008-60 ..
Align seg 1/1 to: US-08-476-008-60 from: 1 to: 427

67 GGCCGGCCATTACGAGTCGCGTTG.....CTTGGCGAAGATA 107
    |||||
81 GlyLysLeuValAlaGlnGlnProLeuGlnLeuPheLeuGlyAsnAlaG 97
108 TGCCGGTATGGCGCC.....TCGATGAAGTCAAGAAAGCGGATG 148
    |||||
97 yThrAlaMetArgProLeuAlaAlaLeuCysLeuGlyLysAsnAspL 114
149 CCGTCAAAAAAGCCCAAGTCTGTTGAGACAAAGAAATCCGGCGCTG 198
    |||||
114 leValLeuThrGlyGlnProArgMetLysGluArg..... 125
199 GTGTTTACTGCGCGGCTTCAGCGAAGAAATGCGCGCATTCACGCTGCGA 248

```


126ProIleGlyHisIleuValaspIalaIleuArgIleGly 138
 249 AAGCGCTACTTACGTAGTCGTGATTCGGTTGAAGCAGCAGCGAA 298
 138 ValaIleIle.....AspIle 144
 299 TCGAGTTGACGCTACGACCTGAGCGGTGGCAACTTAAGCGCGAA 348
 144 euGIleGIleAsnIleTyr..... 149
 349 GAAGTGGCCCGCAACTGATCCAAATCCGTTTGAGCTGCGTGGCAG 398
 150ArgArgCysIleAlaGlyGlyPheArgGlyGlyIleuIle 163
 399 CCGTCCGTTGACGAAATTCCTGCGCTGATCGCCGAGCCGTCGCAAT 448
 163 T.....ValAspGlySerValSerSerGlnP 172
 449 TCGTCAAATGCG...ATGACACCAATCCGCTGCTGCGCCCTACGTC 495
 172 HeleuThrAlaIleuIleuMetThrAlaProIleuIleGlnAspIle 188
 496 ATATCAAGAGCGCCGCGAGGATTTCAACGCGCTGTTGATTTAG 545
 189 IleGlnIleGln.....GlyIleuValSerIle 198
 546 CCGTTTGACCGCAACGCAAAATTCATGTTTGAAGAGCTGGCGCAG 595
 198 SProIleAspIleThrIleuIleuMetIleuIleAspIleValIle 215
 596 TGGCGTGTGAATGCTGCCACATGCAACATGATTCGCGCGCGCG 645
 215 AlValHisGlnAsnIleGlnIlePhe.....HisIleGlyGly... 228
 646 CATCTCGCGCTTGAAGTGCAGCGCATTTCAATCAGCGCGCTGG 695
 229GlnThrTyrArgSerProGlyIleIleValGln..... 240
 696 CCGGATTAACCGTGTGACCATCAATTAATCAATTAATCAATG 745
 241GlyAspIleSerIleAs 247
 746 GCGCTTGTTCACACAGCGCGCTG...AACACGACCGCGTCAAT 789
 247 eTyrPheIleuAlaIleAlaIleIleGlyGlyThrValArgValThr 263
 790 GCGCTAGGTGTTCTCAAGTCAACAAACGCGCTGCTGCTGCTTT 839
 264 GlyIleGlyIleGlnSerValGlnIleAspThrIleAspValIle 280
 840 G.....GGTGGCAAGTATCG.....CAATTA 862
 280 uGIleuIleMetGlyAlaIleIleSerIlePglAspIleIleGly 297
 863 CTGCGGGCGCAATG.....GTTGACACAGCAGCGCGATTCGCGT 906
 297 eArgGlyIleuGlnIleGlnIleAspMetIleAsnIleIleProAsp 313
 907 TCGGATTCGACGCGCGCATTAACAAAGCGCGCATTTATTTGGAGC 956
 314 AlaAlaMetThrIleAlaIleThr..... 321
 957 CTACACAAATCAGATTCCGTTATCGAAGAGCGCGCAGC.....AAG 1000
 322AlaIleuPheAlaAspGlyProThrValIleArg 333
 1001 AGCTGTTCGCTGGTGGCGCGCGCAGCAATTAATCAATCAAGCG 1050
 333 snIleIleTyrAsnIlePargValIleGlnIleAspIleGlnIleMetAla 349
 1051 ACAACCTCGCGCATTTCTGAAAAACAACCTTCAAGTTCACACAGC 1100

350 Thr.....GluIleuArgIleValIleAlaG 358
 1101 CGTCACGCGCGCGCACCGCCCATGATGCCAATGTGACTTACAGCGCG 1150
 358 ValGlnIleuIleGln.....AspIleIleArg 368
 1151 TGATGCCCTTGATATCTGCGCCACCGCTGTTGGCGCATTAATC... 1197
 368 AlValPro.....ProIleGlnIleuIleAlaIleGlnIleGly 380
 1198 ...CTCGCGCATTCGACGACGCGCAGCATTTGCTTGGTGAATGCA 1244
 381 ThrTyrAsnAspHisArgMetAlaMetCysPheSerIleuValAlaIleu 397
 1245 CGAAGAACCTGCGCTTGTGACGCTTCGTCGCCGCGCAAA 1287
 397 rAspThrProValThrIleuAspProIleuAspIleAlaIle 411

seq_name: /cgn2.6/plodata/1/1aa/5A_COMB.pep:US-08-306-063-60
 seq_documentation_block:
 ; Sequence 60, Application US/08306063
 ; Patent No. 5633435
 ; GENERAL INFORMATION:
 ; APPLICANT: Barry, Gerard F.
 ; APPLICANT: Kishore, Ganesh M.
 ; APPLICANT: Padgett, Stephen R.
 ; APPLICANT: Stallings, William C.
 ; TITLE OF INVENTION: Glycosylase Tolerant
 ; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-phosphate synthases
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
 ; STREET: 700 Chesterfield Village Parkway
 ; CITY: St. Louis
 ; STATE: Missouri
 ; COUNTRY: USA
 ; ZIP: 63198
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/306,063
 ; FILING DATE: 13-SEP-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,611
 ; FILING DATE: 28-AUG-1991
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/576,537
 ; FILING DATE: 31-AUG-1990
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hoerner Jr., Dennis R.
 ; REGISTRATION NUMBER: 30,914
 ; REFERENCE/DOCKET NUMBER: 38-21(10660)A
 ; TELEPHONE: (314)537-6047
 ; TELEFAX: (314)537-6047
 ; INFORMATION FOR SEQ ID NO: 60:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 427 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-306-063-60
 alignment_scores:
 Quality: 92.00 Length: 431

Ratio: 0.474 Gaps: 21
Percent Similarity: 45.012 Percent Identity: 20.882

alignment block:
US-09-303-518d-125 x US-08-306-063-60 ..

Align seg 1/1 to: US-08-306-063-60 from: 1 to: 427

```

67 GGGCGGCGCATACCGAAGTCCGCTG.....CTTGGCGAAGATA 107
   ||| .....||| ||| .....||| .....
81 GLLysLeuValAlaGlnProLeuGluLeuPheLeuGlyAsnAlaG1 97
   ||| .....||| .....||| .....||| .....
108 TGGCGGTATGCGCCCC.....TCGATGAAGTCAAGGAGCGCATG 148
   ||| .....||| .....||| .....||| .....
97 yThrAlaMetArgProLeuAlaAlaLeuGlyLeuGlyAsnAspI 114
   ||| .....||| .....||| .....||| .....
149 CCGTCMAAAAGCCCAAGTCGCTTTGAAGACAAAAGATCCGGCGCTG 198
   ||| .....||| .....||| .....||| .....
114 LeValLeuThrGlyGluProArgMetLysGluArg..... 125
   ||| .....||| .....||| .....||| .....
199 GTGTTACTGCGCGGCTTCAGGCAAAATCGCGGATTCACCGTGGCGA 248
   ||| .....||| .....||| .....||| .....
126 .....ProIleGlyHisLeuValAspAlaLeuArgGlnGly 138
   ||| .....||| .....||| .....||| .....
249 AAGCGCGTACTTCAGTCAAGTCGTGATTCGCGTGAAGGCAAGCGAAA 298
   ||| .....||| .....||| .....||| .....
138 yAlaGlnIle.....AspTyrFL 144
   ||| .....||| .....||| .....||| .....
299 TCGAGTTGAAGCTACGACACGTGAAGCGCTGGCAAACTTAAGCGCGAA 348
   ||| .....||| .....||| .....||| .....
144 euGlnGlnGluAsnTyr..... 149
   ||| .....||| .....||| .....||| .....
349 GAAATGCGCGCAACATGATTCATCCGTTGTGACTGCGCGTGGCGAC 398
   ||| .....||| .....||| .....||| .....
150 .....ArgArgCysIleAlaGlyGlyPheArgGlyGlyLeuThr 163
   ||| .....||| .....||| .....||| .....
399 CCGTCGCTGACGAAATTCCTGCGCTGATGCGGAGCGCTTCGCGATCT 448
   ||| .....||| .....||| .....||| .....
163 T.....ValAspGlySerValSerGlnP 172
   ||| .....||| .....||| .....||| .....
449 TCGTCATGCG.....ATGACACCAATCCGCTGCGCGACCTACGGTC 495
   ||| .....||| .....||| .....||| .....
172 heLeuThrAlaLeuLeuMetThrAlaProLeuAlaGlnGlnAspThrGln 188
   ||| .....||| .....||| .....||| .....
496 ATATCAAAAGAGCGCGCGAGGATTCAAACGCGCGCTGTGATTTAG 545
   ||| .....||| .....||| .....||| .....
189 IleGlnIleGln.....GlyGluLeuValSerI 198
   ||| .....||| .....||| .....||| .....
546 CCGTTTGAACGCAACCAATCCATGTTGTGAAGCGATGCGCGACAGC 595
   ||| .....||| .....||| .....||| .....
198 sProtyrIleAspIleThrLeuHisLeuMetLysAlaPheGlyValAspV 215
   ||| .....||| .....||| .....||| .....
596 TGGCGTCTGAATTCCTGCCAAATCGAAACACATGAATTCGGGGCGCG 645
   ||| .....||| .....||| .....||| .....
215 aValHisGlnAsnTyrGlnIlePhe.....HisIleGlyGly 228
   ||| .....||| .....||| .....||| .....
646 CATCCCTGCCGCTTGAATGCGACACACATTCATTCATTCAGCGCGTGG 695
   ||| .....||| .....||| .....||| .....
229 .....GlnThrTyrArgSerProGlyIleTyrLeuValGln..... 240
   ||| .....||| .....||| .....||| .....
696 CCGCAATAAACCGTGTGACACATCAATATCAAGATGATATTCACATG 745
   ||| .....||| .....||| .....||| .....
241 .....GlyAspAlaSerSerIAs 247
   ||| .....||| .....||| .....||| .....
746 GCGGTTGTGTTGCAACAGCGCGCTG.....AACACGAGCGCGTGAAT 789
   ||| .....||| .....||| .....||| .....
247 eTyrPheLeuAlaAlaAlaIleLysGlyGlyThrValArgValThr 263
   ||| .....||| .....||| .....||| .....
790 GCGGTAGGTGTTCAAAGTCAACAAACGCGCGCTTCGCTACCGCTTT 839
   ||| .....||| .....||| .....||| .....
264 GlyIleGlyLysGlnSerValGlnGlyAspThrLysPheAlaAspValLe 280
   ||| .....||| .....||| .....||| .....
840 G.....GGTGGCAAGTATCG.....CAAATTA 862

```

```

280 uGluLysMetCylValAlaLysIleSerITrpgLysAspTyrIleGluCys 297
   ||| .....||| .....||| .....||| .....
863 CTCGCGCGCAATTC.....GTGACACAGACAAACCGCGATTCGCCGT 906
   ||| .....||| .....||| .....||| .....
297 eArgGlyGluLeuGlnGlyIleAspMetAspMetAsnHisIleProAsp 313
   ||| .....||| .....||| .....||| .....
907 TCGGTATGAAACGCGCGATTCACAAAGCGCGCACGATATTTGGAGC 956
   ||| .....||| .....||| .....||| .....
314 AlaAlaMetThrIleAlaThrThr..... 321
   ||| .....||| .....||| .....||| .....
957 CTACCAACATGATGATTCGTTATCGAAGAGCGCGAC.....AAG 1000
   ||| .....||| .....||| .....||| .....
322 .....AlaLeuPheAlaAspGlyProThrValIleArg 333
   ||| .....||| .....||| .....||| .....
1001 ACCTGTTCGGCTGGGTTCGCGCGACCGCGCAAAATCTCCATCAGCGT 1050
   ||| .....||| .....||| .....||| .....
333 snIleTyrAsnTrpArgValLysGluThrAspArgLeuSerAlaMetAla 349
   ||| .....||| .....||| .....||| .....
1051 ACACCCCTGCGCATTCCTCGTGAACAACTCTTCAAGTTCAACACAGC 1100
   ||| .....||| .....||| .....||| .....
350 Thr.....GluLeuArgLysValGlyAlaG1 358
   ||| .....||| .....||| .....||| .....
1101 CGTCAAGCGCGCGACCGCGCATGTGTCGATGCTTACGAGCGCG 1150
   ||| .....||| .....||| .....||| .....
358 uValGlnGlnGlyGln.....AspTyrIleArgV 368
   ||| .....||| .....||| .....||| .....
1151 TGATGCGCTTGATTCCTGCCCACCTGCTTTTGGCGGATTAATC... 1197
   ||| .....||| .....||| .....||| .....
368 aValAlaPro.....ProAlaGlnLeuIleAlaGluIleGly 380
   ||| .....||| .....||| .....||| .....
1198 ....GTGCGCGATACGACAGCGCGCGACGATTCGCTTGAATGGA 1244
   ||| .....||| .....||| .....||| .....
381 ThrTyrAsnAspHisArgMetAlaMetCysPheSerLeuValAlaLeuSe 397
   ||| .....||| .....||| .....||| .....
1245 CGAAGAAAGACCTTCGCTTGTGCAAGCTTCGTCGCCGCGCAAA 1287
   ||| .....||| .....||| .....||| .....
397 rAspThrProValThrIleLeuAspProLysCysThrAlaLys 411
   ||| .....||| .....||| .....||| .....
seq_name: /cgn2.6/ptodata/1/1aa/5A_COMB.dep:US-08-833-485-60
seq documentation block:
; Sequence 60, Application US/08833485
; Patent No 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glycosylate Tolerant
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,485
; FILING DATE: 07-Apr-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 07/749,611
 FILING DATE: 28-AUG-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/576,537
 FILING DATE: 31-AUG-1990
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Hoerner Jr., Dennis R.
 REGISTRATION NUMBER: 30,914
 REFERENCE/DOCKET NUMBER: 38-21(15117)A
 TELEPHONE: (314)737-6047
 TELEFAX: (314)737-6047
 INFORMATION FOR SEQ ID NO: 60:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 427 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-833-485-60

alignment_scores:
 Quality: 92.00 Length: 431
 Ratio: 0.474 Gaps: 21
 Percent Similarity: 45.012 Percent Identity: 20.882

alignment_block:
 US-09-303-518D-125 x US-08-833-485-60 ..

Align seg 1/1 to: US-08-833-485-60 from: 1 to: 427

```

67 GGGCCGGCCATTACCGAAGTCGGTGG.....CTTGGCGAAGATA 107
   ||| :|||: ||| ||| ||| ||| ||| ||| ||| ||| |||
81 GtlyLylLeuValAlaGlnProLeuGluLeuPheLeuGlyAsnAlaG 97
   ||| |||: ||| :|||: ||| :|||: ||| :|||: ||| :|||
108 TGGCGGTATGCGCCC.....TCGATGAAATGCAAGAGAGCGATG 148
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 yThrAlaMetArgProLeuAlaAlaLeuGlyLeuGlyAsnAspI 114
   ||| |||: ||| :|||: ||| :|||: ||| :|||: ||| :|||
149 CCGTCAAAAAGGCCAAGTGGCTTTGAAGACAAAAGATCGGGCGTG 198
   ||| |||: ||| :|||: ||| :|||: ||| :|||: ||| :|||
114 LeValLeuThrGlyGluProArgMetLysGluArg..... 125
199 GTTTACTGCGCGCTTCAAGCAAAATCGCGGATTCACCGTGGCGA 248
   ||| :|||: ||| :|||: ||| :|||: ||| :|||: ||| :|||
126 .....ProIleGlyHisLeuValAspAlaLeuArgGlnGlyG 138
249 AAGCGCTACTTCAATGATGTCGATTCGCGTTGAAGCAAGCGAAGA 298
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 yAlaGlnIle.....AspTyrIle 144
299 TCGAGTTGAAGCGTACGACCTGTAAGCGCGCAAACTTAAGCGCGAA 348
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 euGlnGlnGlnAsnTyr..... 149
349 GAAGTGCAGCCCAACTGATCCAAATCCGTTTGTGAGCTGCGTGGCAG 398
   |||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 .....ArgArgCysIleAlaGlyGlyPheArgGlyGlyLeuTh 163
399 CCGTCCGTTCAAGCAAAATTCCTGCGTCGATGCGGAGCGCTTGCATCT 448
   ||| |||: ||| :|||: ||| :|||: ||| :|||: ||| :|||
163 T.....ValAspGlySerValSerSerGlnP 172
449 TCGTCATATGCG...ATGACACCAATACCGTGGCTGCGCCCTAGAGTC 495
   |||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 heLeuThrAlaLeuLeuMetThrAlaProLeuAlaGlnGlnAspThrGln 188
496 ATTATCAAGAGAGCGCGCGAGATTTCAAAAGCGCGCTGTGTTGATTGAG 545
   ||| :|||: ||| ||| ||| ||| ||| ||| ||| ||| |||
189 IleGlnIleGln.....GlyGlnLeuValSerI 198
  
```

```

546 CCGTTTGACCAAGCCAAATCCATGTTTGTAGGAGCTGGCGGAGAG 595
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
198 sProTyrIleAspIleThrLeuHisLeuMetLysAlaPheGlyValAspV 215
596 TGGCGTGTGAATATGCTGCGACATGCGAACAACATGAAATTCGCGCGCG 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
215 AlValHisGlnAsnTyrGlnIlePhe.....HisLeuGlyGly... 228
646 CATCTCCGCGTTTGTAGTGGACGACATTCATTTGATTCATGAGCGGTCGG 695
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 .....GlnThrTyrArgSerProGlyIleTyrLeuValGln..... 240
696 CGCGAATMAAACCGTGTGACATCAATATGCAAGATGAATATACCATTTG 745
   |||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 .....GlyAspAlaSerSerAla 247
746 GCGGTGTGTTGCAACAGCCCTCTG.....AACCGAGCGCGTGAAT 789
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 eTyrPheLeuAlaAlaAlaAlaIleLysGlyGlyThrValArgValThr 263
790 GCCCTAGTGTGTCTCAAGTCAACAACCGCGCTTGGTACCGTTT 839
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
264 GlyIleGlyLysGlnSerValGlnGlyAspThrLysPheAlaAspValLe 280
840 G.....GGTGGCAAGATATCG.....CAATTA 862
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
280 uGlnLysMetGlyAlaLysIleSerTyrGlyAspAspTyrIleGlyCys 297
863 CTGCGGCGCAATTC.....GTTGACACAGACACCGCGTATTCGCGT 906
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 eArgGlyGlyLeuGlnGlnGlyLeuAspMetLysMetLysHisIleProAsp 913
907 TCGGTATTTGAAGCGCGCATTCACACAGCGCGGACGATTTTGGGACG 956
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 AlaAlaMetThrIleAlaThr..... 321
957 CTACCAACAATCAGATTCGTTATCGAAGAGCGCGAGC.....AAAG 1000
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
322 .....AlaLeuPheAlaAspGlyProThrValIleArg 333
1001 ACCTGTGCGCGTGGTGGCGCGGCGGACCGGCAAAATCTCCATCAGCGGT 1050
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
333 snIleTyrAsnTyrArgValLysGlnThrAspArgLeuSerAlaMetVal 349
1051 ACAACCTTCGCGCAATTCCTGAACAAACAACTCTTCAAGTTCAACAGC 1100
   ||| :|||: ||| :|||: ||| :|||: ||| :|||: ||| :|||
350 Thr.....GlnLeuArgLysValGlyAlaG 358
1101 CGTCAAGCGCGGCGACGCGGCATGTCGCGATTCGACTTACGAGCGCG 1150
   |||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 uValGlnGlnGlnGln.....AspTyrIleArg 368
1151 TGATGCGCTTGTGATATCTGCCACCTGCTTTTGGCGGATTAATC... 1197
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368 AlValPro.....ProIleGlnLeuIleAlaGlnIleGly 380
1198 ...GTGCGCATACCGACAGCGCGGCAAGCAATGAGTTGTTGATTTGGA 1244
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
381 ThrTyrAsnAspHisArgMetAlaMetCysPheSerLeuValAlaLeuSe 397
1245 CGAAGAGAGACCTGCTTGTGCAAGCTTCGTCGCCGCGGCAAA 1287
   :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
397 RasPThrProValThrIleLeuAspProLysCysThrAlaLys 411
seq_name: /cgn2_6/plodata/1/1aa/6B_COMB.pep.us-09-137-440-60
seq_documentation_block:
; Sequence 60, Application US/09137440
; Patent No. 6248876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
  
```



```

      358 uvalgluclgln.....AspTyrIleArgV 368
      1151 TGATGCCCTGGATATCTCCACCCGCTTTGGCGATTAAAC... 1197
      368 alvalp.....ProIaIleuIleAlaIaIaIleuIleGly 380
      1198 ...GTCCGATACCGACAGCGGCGCATTTGGTTGGTGAATTGGA 1244
      381 ThrTyrAsnAspHisArgMetAlaMetCysPheSerLeuValAlaLeuSe 397
      1245 CGAAGAGAGACGCTGCTTTGTGCAGCTTCCTCCCGGCAAA 1287
      397 rAspThrProValThrIleuAspProLysCysThrAlaLys 411

seq_name: /cgn2_6/ptodata/1/laa/6B_COMB.pep:US-09-370-368-4

seq_documentation_block:
; Sequence 4, Application US/09370368
; Patent No. 628932
; GENERAL INFORMATION:
; APPLICANT: Anders Vahne
; TITLE OF INVENTION: PEPTIDES THAT BLOCK VIRAL INFECTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: TRIPR.003A
; CURRENT APPLICATION NUMBER: US/09/370.368
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus Type 2
US-09-370-368-4

alignment_scores:
      Quality: 92.00      Length: 312
      Ratio: 0.648      Gaps: 22
      Percent Similarity: 45.513      Percent Identity: 23.397

alignment_block:
US-09-303-518d-125 x US-09-370-368-4 ..

Align seg 1/1 to: US-09-370-368-4 from: 1 to: 521

516 GGATTTCAAACGGCGCTGTGGTAATGAGCGCTTGCACCGCAACGCAAAA 565
      |||||
173 GlyPheGlnAlaLeuSerGlnGlyCysThrProTyrAspIleAsnGln.M 189
566 TCATGTTTGT.....AAGCAGCTGCG.....GCAGAC 594
      |||
189 etLeuAsnCysValGlyAspHisGlnAlaIaMetGlnIleIleArgGlu 205
595 GTCCCGCTGAAATGCTGCCACATCGAACA...CATGATTCGGCGG 641
      |||||
206 IleIleAsnGlnGlnAlaIaAspTyrAspSerGlnHisProIleProI 222
642 CCGCGATCGCTCGCGTTG.....AGTGCAGCGCACATTCATT 679
      |||||
222 yProLeuProIaIleGlnIleuArgAspProArgGlySerAspIle.... 237
680 TCATGAGCGCGGTCCGCGCAATAAACCGTG.....TGG 714
      |||||
238 .....AlaGlyThrThrSerThrValAspGlnIleGlnIleGlnIle 250
715 ACCATCAATTATCAAGATGAATATTACATTTGGCTTTTTCACACAG 764
      |||||
251 MetTyrArgProGlnAsnProValProValGlyAsnIleTyr..... 264
765 CCGCTGTACACCGAGCGCGTGAATGCGCTAGAGTGTCTCAAGTACACA 814
      |||||
265 .....ArgArgTyrPheIleIleGlnIleGlnIleGlnIleGlnIle 814

```

```

815 AACCGCCCTCTTCGCTA..... 832
277 ArgTyrTyrAsnProThrAsnIleLeuAspIleLysGlnGlyProLysG 293
833 ..CGTTT.....TGGTGGCAAGATATCGCAATATTCGCGGCG... 871
293 uProPheGlnSerTyrValAspArgPheTyrLysSerLeuAlaGlnG 310
872 .....AATTGTTGACACAGACACCGCGTGAATT 901
310 lnThrAspProAlaValLysAsnTyrMetThrGlnThrLeu..... 924
902 CCGGTTCCGATTTGAACGGCGGCGATTCACAGAGCGCGCGCATATT 951
325 .....IleGlnAsnAlaAsnPro..... 330
952 GAGCGCTACCATCATGATTTCCGTTATCGAAGAGCGCGGCGCAAGA 1001
331 ..AspCysLysLeuValLeuLysGlyLeuGly..MetAsnProThrLeu 346
1002 GCTGTCGCGCTGAGTTGCGCGCGCGCGGCA..... 1031
347 GluMetLeuThrAlaCysGlnIleValGlyLysProGlnIleLysAla 363
1032 .....CAATATCTCCATCAGCGGTACACCGCTCGCCCATTTGCT 1070
363 gluMetAlaGlnAlaLeuLysGlnAlaIleMetGlyProSerProIle 380
1071 .....GAATAACAACCTCTCAATTCACAC..... 1097
380 heAlaAlaIaIaGlnIleGlnIleArgLysAlaIleArgTyrTyrPasn 396
1098 .....AGCGTCAACGGCGCGGCGCGCGCGCGCGCGCGCGCGCG 1135
397 GluGlnHisSerAlaArgGlnCysArgAlaProArgArgGlnGlyCys 413
1136 GTACTTACAGCGCGCTGATCCCTTGATATCTCGCCACCGCTGCTTT 1185
413 P.....LysCysGlyLysProGlnHis...IleMet 423
1186 CGCGATTTATGCTGCGGCGATACCGACGCGCGCGCGCGCGCGCGCT 1235
423 laAsnCysProGlnArgGln.....AlaGlyPheLeu 433
1236 GAATTGGACGAGA.....AGACCTCGC 1259
434 GlyLeuGlyProArgGlyLysLysProArg 443

seq_name: /cgn2_6/ptodata/1/laa/5A_COMB.pep:US-08-427-097-2

seq_documentation_block:
; Sequence 2, Application US/08427097
; Patent No. 568294
; GENERAL INFORMATION:
; APPLICANT: Meagher, Richard B.
; APPLICANT: Sommers, Anne O.
; TITLE OF INVENTION: Metal Resistance Sequences and
; TITLE OF INVENTION: Transgenic Plants
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/427,097
 FILING DATE: 21-Apr-1995
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Feiber, Donna M.
 REGISTRATION NUMBER: 33,878
 REFERENCE/DOCKET NUMBER: 40-94
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 499-8089
 TELEFAX: (303) 499-8080
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 564 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-427-097-2

alignment_scores:
 Quality: 92.00 Length: 421
 Ratio: 0.479 Gaps: 17
 Percent Similarity: 45.606 Percent Identity: 20.190

alignment_block:
 US-09-303-518d-125/rev x US-08-427-097-2 ..

Align seg 1/1 to: US-08-427-097-2 from: 1 to: 564

```

1122 GGGCGGTCGCGCGCGCTTGACGCTGTGTAAGTTCGAGTTCGTTT 1073
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 GlyIyIleAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 179
1072 TCAGGAATGCGC.....GAGGTTGTACGCGCGATGAG 1038
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
179 aGlnGlnGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 196
1037 TATTTGTCGCGCGCGCGCTTGACGCGCTTCGTTT.....GCT 994
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
196 Le.....LeuGlnGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 209
993 GCGCGCTTCCTTCGATTAACGGAATCTGATTTGCTGA.....GC 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
210 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 226
955 GTCCCAATATATCGTGGCGCGCTTCGATTCGCGC.....919
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
226 yGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 243
918 .....GTTCAATACCGACCGCA 901
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
243 roAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 259
900 AATCAACGCGGTGTCTGT.....883
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
260 ThrGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 276
883 .....883
276 ySerSerValIeuaIaIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 293
882 .....GTCAACCAATTCGCGCGAGTAATTCGATTCGACACCGCA 838
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293 IalYValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 308
837 AACGCTACCAAGAGCG.....CGGTTGTTGA 809
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
309 ProAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 325
808 CTTGGAACACCACTAGGCAATACGCGCTGCTTCACAGCGCGCTGTT 759
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
325 uValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

758 GCAACAAACGCGCAATGTAATTACATTCGTAATGATGCTCAAC 709
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
338 IeasngIyGlnGlyAspIyGlnIuIaIaIaIaIaIaIaIaIaIaIa 352
708 GGTTCATTCGCGCGCGCGCGCTTCGATTAATGTAATGATGCTCA 659
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
353 GlyIuIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 369
658 AACCGCGCAGG.....ATCGCGCGCGCG 636
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
369 nIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 386
635 AATCATGCTGTTTCGATGTTGCG.....ACATTTTCACAGCGCAC 595
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
386 IYAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 402
594 GTCTGCGCGCGCGCTTCACAAACATGATTCGCTGCGGTCAACGCG 545
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
403 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 415
544 TCATATCAACAGCGCGCTTCGAAATCCGCGCGCTTCGATGATG 495
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
416 .....ValAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 428
494 ACCGTAGGCTCGCGCGCGCGCGATTCGCTGCTC.....462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
429 ThrGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 445
461 .....ATCGATTCAGCAAGATGCGCAACGCGCTGCGCGCATGCA 425
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 eIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 462
424 CGCGAGCAATTTTCGTAACGAGCGGTCGCGCGCATGCTCAACAG 375
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
462 IAspGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 478
374 GATTCGATTCAGGTCGCGCGCGCGCGCTTCGCGCGCTTCGATG 325
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
479 ArgAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 495
324 TTCAGTGCCTGACGCTTCAAACTGATTCGTCGCTTCCTTCAACG 275
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
495 uGlnGlySerGlyArgIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 505
274 TCACGACTGACTGACGATGACGCGCTTCGCGCGCGATTCGCGG 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
506 .....AlaVal 507
224 TTGCTGAAGCGCGCGCGAGTAACACCGCGCGCGATTCCTTTGCT 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
508 AlaProGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 524
174 AACAGCACTTCGCGCTTCGACGCGATC.....CCTCTCTGACTT 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
524 nArgMetThrValGlnIuIeuaIaIaIaIaIaIaIaIaIaIaIaIa 541
130 TCATCGAGCGG 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
541 eValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa

```

seq_name: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:US-08-427-097-14
 seq_documentation_block:
 ; Sequence 14, Application US/08427097
 ; Patent No. 5668294
 ; GENERAL INFORMATION:
 ; APPLICANT: Meagher, Richard B.
 ; APPLICANT: Sommers, Anne O.
 ; TITLE OF INVENTION: Metal Resistance Sequences and
 ; TITLE OF INVENTION: Transgenic Plants
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, winner and Sullivan, P.C.

seq_name: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:US-08-427-097-16

```

seq documentation block:
Sequence 16, Application US/08427097
Patent No. 5668294
GENERAL INFORMATION:
APPLICANT: Meagher, Richard B.
APPLICANT: Sommers, Anne O.
TITLE OF INVENTION: Metal Resistance Sequences and
TITLE OF INVENTION: Transgenic Plants
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,097
FILING DATE: 21-APR-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Feher, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 40-94
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-427-097-16

alignment_scores:
Quality: 92.00 Length: 421
Ratio: 0.479 Gaps: 17
Percent Similarity: 45.606 Percent Identity: 20.190

alignment_block:
US-09-303-518d-125/rev x US-08-427-097-16 ..
Align seg 1/1 to: US-08-427-097-16 from: 1 to: 564

1122 GCGCGGTCGCGCGCGCTGACGCGCTGTGTAAGTTGTTT 1073
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 GYGLYILEALALALALALALALALALALALALALALAL 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1072 TCAGCAATGAGCC.....GAGGTTGTACGCGTGATGAG 1038
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 agtnglgnlgnlalnargvalasplulemarghisalalystyrgl 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1037 TATTTTCGCGCTCGCGCCCAACCGCAACAGCTCTT.....GCT 994
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 le.....leuglgnlgnasnproalalalethvalleuhiselyser 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
993 GCGCGCTTTCGATACGAAATCGATTGTGTA.....GC 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 Alalnrgphelysaspasnargasnleuilevalgnleuasnasp 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
955 GTCCCAATATATCGTCGCGCTGTGTATATCGCGC..... 919
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
226 ygluargvalvalalaphaspargysleuilealathrglala 243
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
918 .....GTTCAATACGCAACCGCA 901
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

243 roalavalprolileproglileuylsaspnrhprolytrptlrhser 259
900 AATCAGCGGTTGTCTGT..... 883
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
260 Thrghlualaleuvalsergluthrileprolysargleualalel 276
883 ..... 883
276 yserSerValleuAlaleuGluleuAlaglnAlaphAlaArgleuGly 293
882 .....GTCAACCAATTCGCCCGCAGTAAATTCGCTACTTCGACCA 838
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 lalysValthrilleuAlaArgSerThrleu...PhepheArgleuasp 308
837 AACGTCACGAGAGCGG.....CGGTTGTGTA 809
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
309 prolaAllelglylvalalvalAlalapharpmetgluglylle 325
808 CTTCAGAACCCACTAGCGAATCAGCGCTCGTTCAGACGCGCTGT 759
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 uvalArg.....GluHisThrGlnAlaserGlnvalAlaIlyr 338
758 GCAACAAACGCGCAATGTATATACATCTGATATATGATGTCACAC 709
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
338 leasnnglygluglyaspBglylnphevalleuThr.....Alahis 352
708 GGTATTATTCGCGCGCGCGCTCGATGAATGAATGTCGTCACATCA 659
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 GlyluleuargAlaaspysleuValAlalathrglyAlaAlapros 369
658 AACCGGACG.....ATCGCGGCGCGCG 636
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 nThrArglysleuAlaleuaspAlaThrlyAlthrleuThrproGln 386
635 AATCAGTGTTCGATGTTGCG.....AGCATTTTCAGACGCGCAC 595
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 lyalAllevalleAlasprgylmetArghrSerValGluHisIleYr 402
594 GTCTGCGCGACGCTCTTACAAACATGATTTTCGTCGGTCAACGCGC 545
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
403 AlaAlaglyaspysrhnAspGlnProGlnpheValIlyr..... 415
544 TCATATACCAACAGCGCGGTTGAAATCTCGCGGCTCTTGTGATATG 495
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 .....ValAlaAla..AlaAlaglyThrhrAlaAlaIleasmet 428
494 ACCGTAGGTCGCGACCGCGGATTGTCTC..... 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
429 ThrGlyglaspAlaAlaleuasnleuThrAlaserProAlaValAlph 445
461 .....ATCGCATTCAGCAAGATGCGCAACGCTCGCATCGA 425
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
445 eThrAspProGlnValAlaThrValAlgIlyrSerGlnAlaGlnAlahis 462
424 CGCAGCAATTTTGTGTAACGAGCGGTGCGACGCAAGTCACAAACG 375
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
462 lAspBglylIleYrThrAspserArghrThrleuAlaAspAsValpro 478
374 GATTGATCAGCTTCGCGCGACCTTCGCGGCTTAAGTTTGCCACGCG 325
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
479 ArgAlaAlaAlaAsnPhAspThrArglylphelleYleuValAlgl 495
324 TTCAGGTGCGTAGCTTCACAACTCGATTTCGTCGCTTCACAGCGCA 275
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
495 ueluglyserGlyArgleuIleGlyValGln..... 505
274 TCACGACTGACTGAATACGCGCTTTTCGCCACGCTGAATGCGCGGATT 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 .....AlaVal 507
224 TTCCTGAGCGCGCGAGTAACACACGCGCGGATTTCTTTGTCTTC 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
508 AlaproglyAlaglyluleuileGlnThrAlaAlaIleAlaIleargAs 524

```



```

324 TTCAGGTGCGTACGCTCAAACTGATTTCGTGCTTCAACGCAA 275
      ::|||::: |||
      ::|||::: |||
495 uGluGlySerGlyArgLeuIleGlyValGln..... 505
      ::|||::: |||
274 TCACGACTGACTGAAGTACGGCGCTTTCGCCACGGTGAATCGGGCGATT 225
      ::|||::: |||
506 .....AlaVal 507
224 TTGCCTGAAGCGCGCGCAGTAACACACACGCCCGGATTCTTTTGTCTTC 175
      ||||| |||||
508 AlaProGlnAlaGlyGlnLeuIleGlnThrAlaLeuAlaIleArgAs 524
      ||||| |||||
174 AAACAGCACTGGCCTTTTGTACGGCATCG.....CCTTCCTTGACTT 131
      ::|||::: |||
524 nArgMetThrValGlnGlnLeuAlaAspGlnLeuPheProTyrLeuThrM 541
      ::|||::: |||
130 TCATCGAGGGG 120
      ::|||::: |||
541 etValGlnGly 544

```